Cross Platform Web-based Smart Tourism Using Deep Monument Mining

Mehdi Etaati
Department of Computer Engineering
Khatam University
Tehran, Iran
m.etaati@khatam.ac.ir

Babak Majidi
Department of Computer Engineering
Khatam University
Tehran, Iran
b.majidi@khatam.ac.ir

Mohammad Taghi Manzuri
Department of Computer Engineering
Sharif University of Technology
Tehran, Iran
manzuri@sharif.edu

***Abstract*— Tourism is one of the largest sources of economic revenue for many countries around the world. The historical and cultural treasures of Iran made it one the main destinations for international tourists. One of the biggest problems encountered by the tourists during the visit to monuments of Iran is the lack of information about the visited landmark. Given that cameras can be found in all of the smart phones, the use of the landmark's photos can be very important for obtaining information about the tourism sites. The detection of the landmarks in an image taken by the mobile phone camera can be a very complex task depending on the angle and the light situation in which the photo is taken. In this paper, a web based cross platform mobile framework based on deep neural networks for autonomous identification of the historical landmarks of Iran is presented. The images recorded by the mobile phone is sent to the decentralized servers in order to be processed and the information about the landmark is determined and transferred to the mobile device of the tourist. The proposed framework is evaluated on the tourism attractions of Iran and the experimental results show that the proposed system can recognize the historical landmarks with precision of 95%.**

Index Terms*— *smart tourism, web-based applications, convolutional neural networks, deep learning.

I. INTRODUCTION

Due to its ancient culture, art and architecture, Iran attracts a large number of tourists from around the world every year. A large number of dazzling monuments with unique architectures in Iran, made it a must-see destination for anyone interested in the art and culture of the world. In every city of Iran there is a large number of historical sites, each with unique historical and cultural story behind it. Due to large number of these landmarks in Iran, an autonomous and intelligent system for providing information about these monuments is necessary. Currently, all tourists have a mobile phone camera and utilizing the visual information is a solution to avoid the language barriers when tourists require information about a landmark. The large database of images of these landmarks on internet make the images the optimal solution for information retrieval without knowing the name of the landmark.

Computer vision and machine learning techniques are able to provide reliable intelligent autonomous solutions for many applications [1-4]. Intelligent tourism applications are gradually introduced into the tourism industry and there is a significant recent surge in research about these applications. Zaifri et al. [5] proposed a smart phone based augmented reality for smart tourism. Sergeeva et al. [6] presented a model for 3D reconstruction of landmarks from the images. Noinan et al. [7] developed a smart tourism app for mobile phones. This app recommends a tourism route to the tourist in Thailand. Maach et al. [8] developed a virtual reality system for visiting the historical landmarks. Demir et al. [9] proposed an augmented reality application for smart tourism. Dangkham [10] developed a web based augmented reality platform for tourism using HTML5 standards. Butgereit et al. [11] presented a smart tourism platform for areas in Africa. Wei et al. [12] investigated the application of real time processing of tourism based big data. Wanchun [13] proposed a tourist assistant mobile based platform. Shukla et al. [14] presented a computer vision framework for autonomous description of tourism landmarks in India.

Deep neural networks are recently used for solving complex computer vision and pattern recognition tasks [15-17]. In this paper, a web based mobile framework based on a deep neural network for automatic recognition of the historical landmarks of Iran is presented. A deep convolutional neural network is used for training a model of tourism attractions of Iran. The smart phone videos are processed on decentralized servers to identify the landmark for smart tourism. The proposed framework is evaluated on the historical landmarks of Iran. The experimental results show that the proposed framework can identify the tourism landmarks with precision of 95%.

The rest of this paper is organized as follows. The proposed cross platform web-based smart tourism using deep monument mining framework is detailed in Section II. The experimental design and the simulation scenarios are discussed in Section III. Finally, Section IV concludes the paper.

II. CROSS PLATFORM WEB-BASED SMART TOURISM USING DEEP MONUMENT MINING

The proposed system architecture is completely decentralized and platform independent. The architecture of the proposed framework is presented in Fig. 1. The user, records the image of the landmark from any perceptive using the smart phone camera.

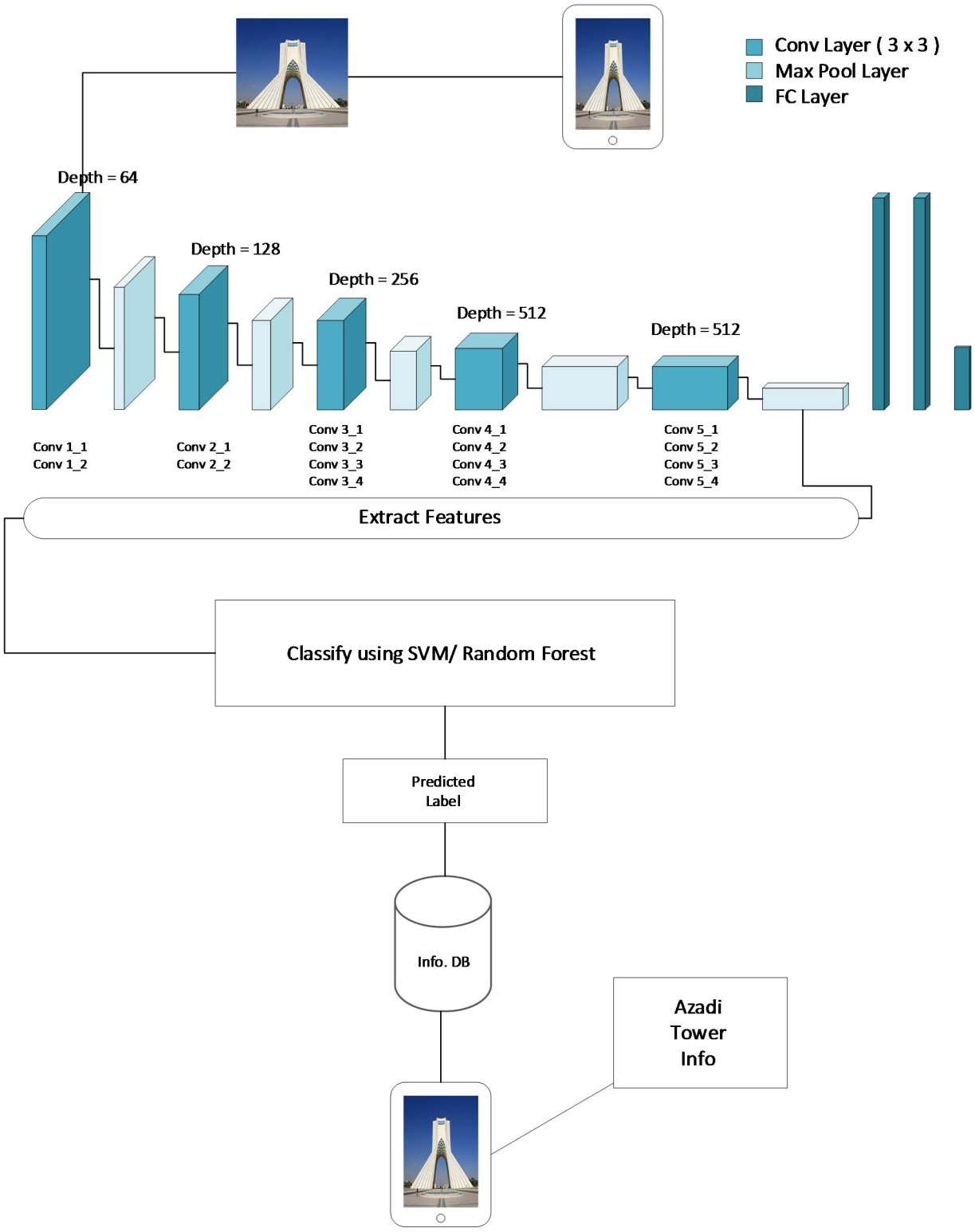


Fig. 1. Deep monument mining architecture

The images are sent to the nearest web-server based on the tourist location and IP. The server processes the images and, after detecting the monument found in the image, the building information from the database extracted and the results are sent to the tourist's device. Due to the complexity of the architectures of monuments in Iran and resemblances of various historical sites to each other, it is very important to provide a solution with high accuracy to identify the landmarks.

The task of the monument detection sub-system is to identify the buildings and to separate them from the background objects in the image. Extraction of discriminating features are the main component of the visual processing system. The use of low-level visual features and local descriptors has been used in many applications for the recognition and classification of buildings in the image. In this research, first the application of the Scale Invariant Feature Transform (SIFT), Speeded Robust Feature (SURF) and Oriented FAST and Rotated BRIEF (ORB) for monument recognition is investigated. The experimental results show the lack of accuracy using these features for complex scenes classification. Another method for complex scene classification is deep neural networks which are used to extract high-level visual features.

For extraction of the monument's features, in this research a deep neural network and transfer learning technique is used. Then the classification algorithms Support Vector Machine (SVM) and Random Forest are used for classification of objects in the image using the extracted features by the deep neural network. The architecture of the deep neural network used in this research is the VGGNet [18] model with 19 layers. The VGG 19 architecture is shown in Fig. 1. The image is given as input to the VGGNet network, then the output of the last MaxPool layer is extracted as the feature vector corresponding to the input image. This vectors dimension is 7 by 7 by 512 which is then flattened to a 1 by 25088 feature vector. In the classification stage a multi-class SVM with rbf kernel and one-to-one strategy and a random forest with 10 estimators are used as classifiers.

III. Experimental Results

In this research a dataset of major Iranian historical landmarks is created for transfer learning using the deep neural network. In order to improve the proposed method, deep neural networks with deeper layers such as RESNET network architecture with 50 layers can be used. Images recorded by tourists are usually contain a large amount of noise. The object detection process in presence of noise can be greatly improved by using the salient object detection method. Also, due to the limited number of images used in this study, by collecting more images in order to improve the transfer learning the model can be improved.

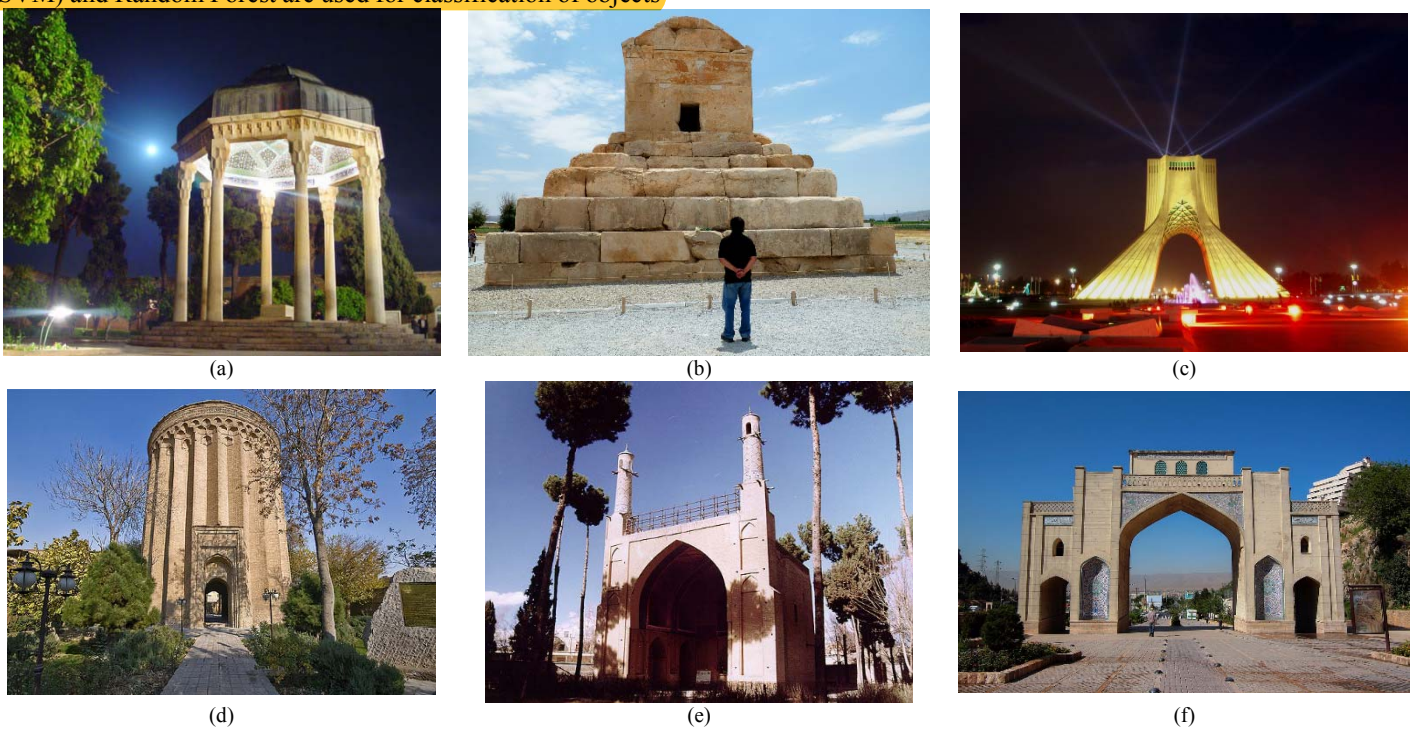


Fig. 2. Samples of Iranian monuments used for training the deep neural network

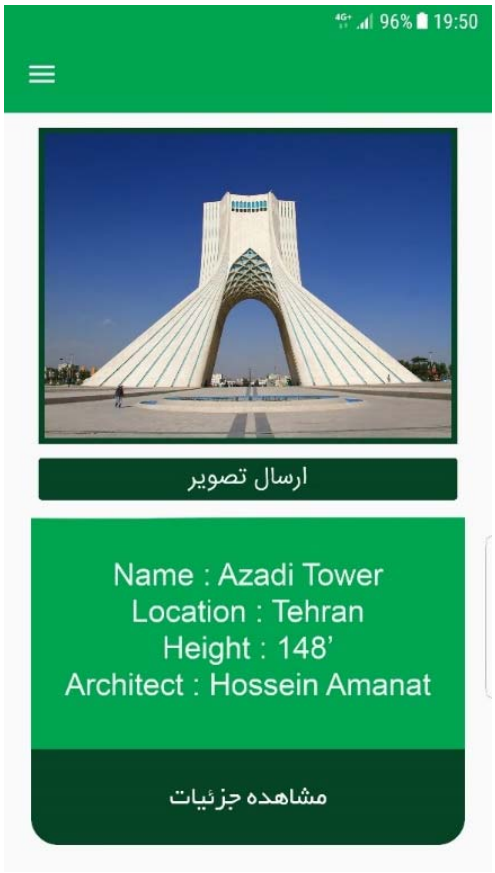
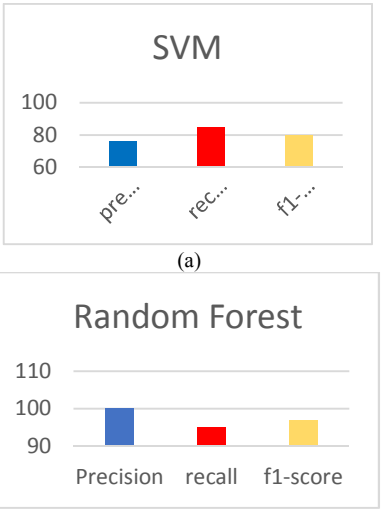


Fig. 3. Cross platform web-based trouism app

Fig 3. Shows the cross-platform web-based app developed for providing the tourist with the information about the monuments. Fig. 4 shows the classification results using SVM and random forest algorithms. In the first step, using SVM, the classification was performed and the accuracy of 85% was achieved. In the second step, using random forest, the accuracy increased to 95%.



(a)

(b)

Fig. 4. Classification results

CONCLUSION

In this paper, a web based cross platform framework based on a deep CNN for autonomous identification of the historical landmarks of Iran for smart tourism is presented. The proposed platform uses a deep neural network for feature extraction from the images captured on a mobile phone and a classification algorithm for identification of monuments in the image. 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 precision of 95%.

REFERENCES

[1] B. Majidi and A. Bab-Hadiashar, "Aerial tracking of elongated objects in rural environments," *Machine Vision and Applications,* vol. 20, pp. 23-34, 2009.

[2] B. Majidi, J. C. Patra, and J. Zheng, "Modular interpretation of low altitude aerial images of non-urban environment," *Digital Signal Processing,* vol. 26, pp. 127-141, 2014.

[3] A. Shamisa, B. Majidi, and J. C. Patra, "Sliding-Window-Based Real-Time Model Order Reduction for Stability Prediction in Smart Grid," *IEEE Transactions on Power Systems,* vol. 34, pp. 326-337, 2019.

[4] A. Mansouri, B. Majidi, and A. Shamisa, "Metaheuristic Neural Networks for Anomaly Recognition in Industrial Sensor Networks with Packet Latency and Jitter for Smart Infrastructures," *International Journal of Computers and Applications,* 2019.

[5] M. Zaifri, A. Azough, and S. O. E. Alaoui, "Experimentation of visual augmented reality for visiting the historical monuments of the medina of Fez," in *2018 ISCV*, 2018.

[6] A. D. Sergeeva and V. A. Sablina, "Using structure from motion for monument 3D reconstruction from images with heterogeneous background," in *2018 MECO*, 2018.

[7] K. Noinan, N. Somprasitwit, S. Jaisungnern, A. Sranoi, K. Nount, and S. Wicha, "Route recommendation for local tourism: The empirical system of Chiang Rai province," in *2018 ICDAMT*, 2018.

[8] I. Maach, A. Azough, and M. Meknassi, "Development of a use case for virtual reality to visit a historical monument," in *2018 ISCV*, 2018.

[9] Ö. F. Demir and E. Karaarslan, "Augmented reality application for smart tourism: GökovAR," in *2018 ICSG*, 2018.

[10] P. Dangkham, "Mobile augmented reality on web-based for the tourism using HTML5," in *2018 ICOIN*, 2018.

[11] L. Butgereit and L. Martinus, "Assisting Tourism in Underserved Areas with TensorFlow: A Proof-of-Concept Mobile App," in *2018 IST-Africa Week Conference (IST-Africa)*, 2018, pp. Page 1 of 7-Page 7 of 7.

[12] J. Wei, L. Ma, and Z. Zhang, "A research on smart tourism-oriented big data real-time processing technology," in *2017 29th Chinese Control And Decision Conference (CCDC)*, 2017, pp. 1848-1851.

[13] Z. Wanchun, "Design and Implementation of Intelligent Tourism Platform for Mobile Terminal," in *2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, 2017, pp. 705-708.

[14] P. Shukla, B. Rautela, and A. Mittal, "A Computer Vision Framework for Automatic Description of Indian Monuments," in *2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, 2017, pp. 116-122.

[15] M. H. Abbasi, B. Majidi, and M. T. Manzuri, "Deep cross altitude visual interpretation for service robotic agents in smart city," in *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, 2018, pp. 79-82.

[16] M. H. Abbasi, B. Majidi, and M. T. Manzuri, "Glimpse-gaze deep vision for Modular Rapidly Deployable Decision Support Agent in smart jungle," in *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, 2018, pp. 75-78.

[17] S. Sanaei, B. Majidi, and E. Akhtarkavan, "Deep Multisensor Dashboard for Composition Layer of Web of Things in the Smart City," in *2018 9th International Symposium on Telecommunications (IST)*, 2018, pp. 211-215.

[18] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 39, pp. 640-651, 2017.