2020 IEEE REGION 10 CONFERENCE (TENCON)
Osaka, Japan, November 16-19, 2020

Historical Places & Monuments Identification System

Navod Godewithana[1], Khema Jayasena[2], Chamodi Nagarawaththa[3], Praveenth Croos[4], Buddika Harshanath[5], Jesuthasan Alosius[6]
Department of Software Engineering[1, 5], *Department of Information Technology*[2, 3, 4, 6]
Faculty of Computing
Sri Lanka Institute of Information Technology
Colombo, Sri Lanka
navodjanitha@gmail.com[1], upsalajayasena@gmail.com[2], chamodirangani@gmail.com[3], praveenthcrooz@gmail.com[4], harshanath.s@sliit.lk[5], jesuthasan.a@sliit.lk[6]

***Abstract*— Sri Lanka, which is known as "the pearl of the Indian ocean" provides great survival and civilization history dating back to the 3rd century. Most of the archaeological sites are attracted by not only Sri Lankans but also by tourists. When searching for the information about the archaeological sites, there are lack of trusted information sources and smart online platforms. Even though some information is available, no convenient and efficient ways to retrieve them. When trusted information is provided in a user-friendly manner, the value will be added to the Sri Lankan economy. Since the world is driving towards the "E-Era", everything is involved with Information Technology. The proposed system contributes to solve the above problems with Artificial Intelligence & Machine Learning concepts. The system is assisted using four major components namely, image identification, community platform, conversational bot, and image visualization. The image Identification component identifies the archaeological sites using image processing techniques. The community platform gathers trusted information from archaeologists and deep learning techniques are used to deliver that content to the users. The artificial intelligence conversational bot is established to communicate and retrieve available information in a convenient manner. The image visualization component is used to provide reality visualization on archaeological sites using the augmented reality techniques. The techniques and the algorithms are evaluated to deliver better performance with brilliant user experience.**

Keywords*— *Artificial Intelligent, Augmented Reality, Chatbot, Machine Learning, Natural Language Processing

I. Introduction

Sri Lanka, known as "the teardrop of India," is a small but amazing island with a great civilization history. Historical, cultural, and religious valued places and monuments are the mirrors of this proud history [1]. All most all the archaeological places are attracted by foreign and local tourists [2]. For an example, when people visit such a place or a monument, with the development of the information age it became a habit to search for information about the place and try to build relationships. But in the Sri Lankan context, there are lack of online resources to find much and accurate information. We critically analyzed the above mention existing resources and platforms to identify the existing problems related to the domain. Even though, there are some existing platforms capable to gain information about those places [3], [4], [5], these platforms are unable to provide accurate information about most of the unpopular historical sites. Furthermore, the available platforms do not provide a convenient way to obtain the available information and visualization mechanisms. The final aim is to develop an innovative solution for this complex problem.

Several research works have been conducted partial work on heritage protection and data management in several countries. As an example, Binyue et al. [6] proposed a system for integrating cultural and heritage data protection into a smart way. Another research work has been discussed the usage of image processing techniques to detect discoloration and deformations in ancient pictures [7]. However, in the Sri Lankan context, new technologies and digital domains have lack of involvement of obvious automatic solution to meet the above fundamental challenges. Tourism is one of a main income generator industry in Sri Lanka and most of the archaeological places are attracted by the tourists [2]. Because of that, the necessity of developing new strategies addressing the usage and management of archaeological data comes into a big picture. The system consists of 4 subsystem modules.

The image identification module is intended to identify a place using an user given image, location-based data or the user inputs in order to accurately identify the archaeological sites. Moreover, the user will obtain information about the places and the relationships it has with other places, people, or sources. The system facilitates a social platform to interact with users and archaeologists who are interesting in this historical domain, where users could contribute their knowledge and information to the system. It maintains the highest accuracy ratio from the given information and uses summarization techniques to maintain the information given by the users. Moreover, the artificial intelligent chatbot provides a smart way of searching the available information on archaeological places. Instead of much more search results, this specialized chatbot narrows down the information for the user while providing a convenient way of gaining the information. Furthermore, using augmented reality techniques, it provides great visualization on Sri Lankan archaeological places as well. The proposed system will use technologies like Deep Learning, Image Processing, Natural Language Processing, Augmented Reality and many other modern and research effective technologies.

II. Literature Review

Given the smart concepts are gaining momentum to provide a solution to sustainable problems. When considering about the existing solutions and the recent research works, there are some online sources available to get information and identify archaeological sites in Sri Lanka. Such as the

UNESCO website [3], Archaeology Travel web service [4], Web site of Sri Lankan Archaeological Department [5] are some of the available online sources that provide information about historical places in Sri Lanka. Even though these websites provide some information about those places, most of the un-popular archaeological sites are unable to identify by the existing sources. And there is a problem with the trustworthiness of the information that users provide. There were unable to find any assemble online resource which could identify images, provide socialization platform, enabling conversational bot for the information retrieval and image visualization for Sri Lankan archaeological sites. When critically analyzing the available resources and recent research works, following main contexts are identified as main areas that should be improved from an innovative system.

- Provide a mechanism to identify archeological places by images
- A community platform that facilitates to interact with users who are interesting in this domain and grow the knowledge base
- AI based chatbot as a smart way of searching and retrieving the available information
- Mechanism to visualize places with 3D content creation using Augmented Reality

Furthermore, when consider about the above four specific component domains, following solutions are identified as existing research works. The paper [8] presents a framework for recognizing monuments images. Monuments recognition is a challenging problem in the domain of image classification due to huge variations in the architecture of different monuments. In here Deep Convolutional Neural Networks (DCNN) is used for extracting representations. The model is trained on representations of different Indian monuments, obtained from cropped images, which exhibit geographic and cultural diversity. These are the important and visual source of analyzing the history of India, very precisely. The manually acquired dataset comprises of 100 folders with each folder having 50 images per monuments. The naming of each folder is done according to the name which corresponds to the monuments. The key intuition behind the monument classification was the use of basic hand-crafted features like histogram of oriented gradients (HOG), Local Binary Pattern(LBP) and Gastrointestinal Stromal Tumor (GIST) to find out whether these features could be used to classify monuments with a better accuracy. CNNs used a variation of multilayer perceptron's designed to require minimal preprocessing. It was able to achieve a much better accuracy of 92.7% in comparison to the hand-crafted features. The experiments performed proved the importance of using representations of monuments images to build an effective monuments recognition system.

The research conducted by Vijay et al. [9] used a hybrid model for summarizing text documents using Text Rank Algorithm and Term Frequency. First, the authors pre-process the given document by removing stop words, tokenize the input text into words, Stemming (Process of removing suffix from words and get the root word), Normalization (Process of transforming the text into the same form such as converting the whole text into lowercase.) and sentence tokenization (Tokenize the text data into sentences). Then generate the Text Rank matrix which will have the ranking of all sentences with the highest score of the most significant sentence. Then calculate the cosine similarity among sentences. After that, build the similarity matrix. Finally, create the summary based on top rank scores from the Text Rank matrix. The outcome is evaluated using the ROUGE metric. The researchers conclude that the hybrid model is best suited for the summarization.

The study conducted by Shankari et al. [10] used the gensim summarizer for multi-document text summarization. Genism is a library which provides by the NLP and it provides the user with a semantic structure by using raw data. Firstly, preprocess the data by removing stop words and convert them to lowercase and use the genism library to summarize the text. Finally, for analyzing, count the number of words in the original text and summarized text and conclude that the genism summarizer is the best for summarization purposes.

A context-aware chatbot was developed by Clarizia et al. [11] for providing cultural heritage information to the tourists. The objective of this research work is much more identical to the our research objective. The researchers identified the importance of cultural heritage and the contribution that can gain with new technologies to enhance the user experience by providing that information using a chatbot to the users. Furthermore, the proposed system is able to recommend contents and services according to tourist profiles and context. The bot is assisted with pattern recognition and context recognition techniques. The proposed architecture includes several main modules with the responsibilities associate with them. The digital story telling module guides user throughout the conversation. Context-Aware Manager is capable of extract and offer personalized information. The first experimental results are satisfactory and show the potential of the proposed approach.

Another chatbot application was invented by Lombardi et al. [12] for cultural heritage in Italy. This invention is proposed to help tourist in the most disparate situations such as getting the right information. The goal of this research work is partially identical to this proposed conversational agent's goal. The proposed methodology uses semantic analyzer, workflow manager, external services and a user interface. Experimental phase has been made into Pompeii archeological park, through a questionnaire given to visitors that have used the app with the Chatbot installed on specific devices connected on the network.

Yapahuwa MR Application for Yapahuwa archaeological site is another application that has Virtual Reality option [13]. The applications like House of obrich [14] platform defendable, since it is only compatible with IOS platform, but it has a good animation to attract users. This application visualizes the history of Darmstadt unique Jugendstil. Heladiva AR [15] is another Augmented Reality application to reconstruct Sri Lankan cultural heritage in Prime State. These authors created a smartphone application to reconstruct historical site. It has detailed modules unique from other similar applications.

III. METHODOLOGY

The system is capable of detecting the archeological places using deep image processing techniques. The community platform provides archaeologists and other users to engage with the platform by providing information on the archaeological sites. It is assisted using deep text summarization techniques to maintain the consistent information base. Furthermore, the platform consists of a conversational agent to gain the available information in a convenient and efficient manner. The image visualization facilitates the virtual visualization on archaeological sites by using augmented reality techniques. The four major components, namely: Image Identification System, Community Platform and maintaining the Information Base, Artificial Intelligent Chatbot and Image Visualization using Augmented Reality. illustrate by the system architecture diagram given in Fig. 1.

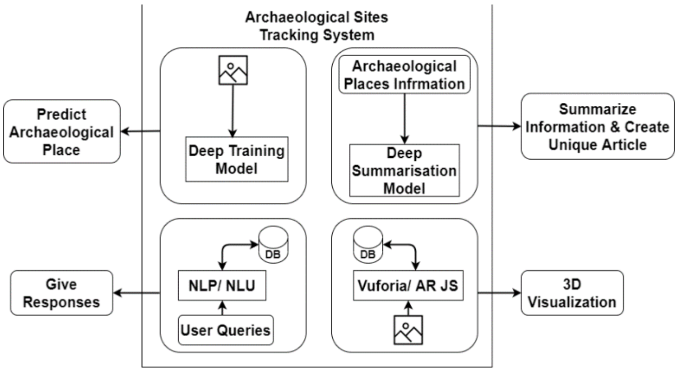

Fig. 1. System architecture diagram

A. Image Identification Module

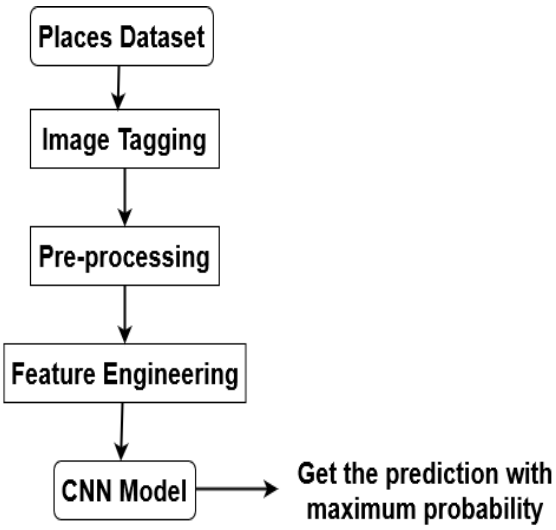

Fig. 2. Image identification process overview

The image identification component is to identify the archaeological sites using the images. This component helps to identify exact location from the image of a place which is uploaded to the system by user or captured by user using a device that is needed for analysis. Fig. 2 shows the architecture of the component. There is a technology called speeded up robust features (SURF) detector, which is used to image detection, feature extraction and learning process of SURF is very different than the processing in the human brain. In the concept of classification, SURF use to object detection, edge detection and also lines, shapes, colors detection but it will not take accurate result. It detects most of the time only shapes in the image. There is currently very popular algorithm for classification, image recognition algorithm which is derived from convolutional neural network (CNN). It is the way of feature extraction and very good generalization from simple to complex is much more similar to the processing in the human visual systems. CNN identifies not only shapes but also whole image and take good accuracy level. According to the process, it creates the data directory with label category using archaeological sites images. Labeled data directory which is gone through the image preprocessing unit because of all images are resized using open cv library. In the feature engineering process, it extracts the features of dataset. Then the data set is gone through layers of convolutional neural network architecture because of model training. These convolutional layers detect features several times from simple to deep and also create features map. The Max pooling layers are done by applying a max filter to keep up essential features and loss unwanted. Then it will make a single image again in flatten. After that, it is resulted outputs with category label according to maximum probability. About 10 times this way, losses are decrease and the accuracy level is increased. Finally save the training model to use for predictions. According to this way it can be predicted archaeological sites using images with maximum probability.

B. Community Platform

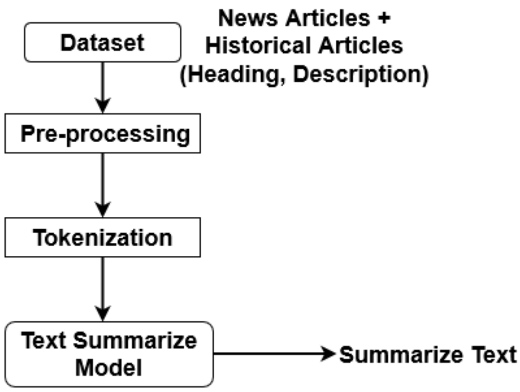

Fig. 3. Text summarization process overview

Users need to spend a lot of time for gathering information about historical places in the current domain. The purpose of this component is to generate summaries to the articles given to the system or published by the system users. From those summaries, the users can obtain an overall idea in the content. The archeologists are the information providers for this system. The description provided by those users save in the database as a text file. When community users request information, the system will display all generated summaries for a given search alone with a hyperlink to access the complete article. Machine learning and deep learning approaches are used for summarization as shown in the Fig. 3 component architecture diagram.

Dataset creation: To train the model, used BBC news article dataset [16] along with the historical articles of Sri Lankan Context, which consists of headings and descriptions.

Data pre-processing: Data is cleaned to remove unwanted characters like punctuation marks and spaces from the text. It helps to remove unnecessary bumps that may come during the training.

Tokenization: Build the dictionary by splitting text into smaller tokens. That helps to filter out unnecessary data.

POS tagging: NLTK package includes a method called "pos_tag" to perform part-of-speech (POS) tagging. It will return only the nouns and verbs from the text.

Model selection: Gensim and term frequency algorithms are used to create extractive summaries. When testing the generated summaries are not accurate as it contains many grammar mistakes and unable to provide an accurate summary for a given article. Since that, used sequence to sequence recurrent neural network algorithm, which is an abstractive technique to generate summaries. The model is trained using 1960 records and tested on the remaining records. Build the model by stacking 3 LSTMs (Long Short-Term Memory) vertically with the horizontal axis. After 30 epochs, the losses decreased, and the accuracy level increased. The seq2Seq algorithm provides a better summary than the other two in this approach.

Deploy the model: The model is deployed on the flask by creating an API. The system uses that API to generate summaries for articles related to a given search.

C. Conversational Agent (Chatbot)

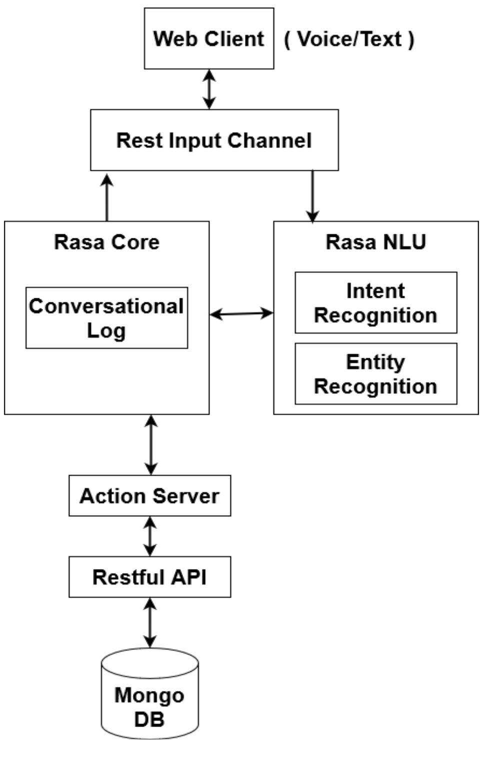

Fig. 4. Chatbot component overview

Rather than normal search mechanisms, a chatbot is a new customer interactive way in most platforms. As shown in the above Fig. 4, User Interface, Natural Language Understanding, Dialog Management (Rasa Core) and the Restful API with Mongo DB are the four main modules assisted by the hybrid chatbot.

The client interaction is established through the user interface which would be recognized both voice and text inputs as user messages, where user could be able to ask questions relevant to the archaeological sites in Sri Lanka in text or voice formats. After that, the user query is passed to the Natural Language Processing module. Natural Language Processing is one of theoretically advanced techniques for the automatic understand human beings and representation of their language [17].

Rasa NLU and Rasa Core is used as a framework to build the NLP module, which is an open source natural language understanding framework [18]. After receiving a user message to the NLU component, it captures the contextual information. It performs tokenization, featurization, entity recognition and intent classification using the received message. WhiteSpaceTokenizer is used for tokenization where it considers white spaces as a separator. The intent classification is implemented using TensorFlow embedding (EmbeddingIntentClassifier), which is trained using domain vocabulary including a one-hot encoded vector (text features) and the results of given regular expression patterns. The above-mentioned domain vocabulary was produced using CountVectorizer. As the output of the Embedding Intent Classifier, it provides recognized intents and the associated intent ranking as the confidence. The CRFEntityExtractor outputs the probability of entities given the sequence. Finally, the EntitySynonymMapper is utilized to map synonyms entity values to the same value. Afterall, the NLU module returns a structured format of the user's message.

Then the intents are passed to the dialog management module which is using Rasa Core approach. It obtains the next possible best action for the user query by using several policies. The policy decides what should be the next action by considering the inputs and its implementation. It uses customized Keras policy, Memorization policy and Fallback policy for the policy stack of this chatbot. The Action server invokes the Rest API backend when it needs to gain data stored in the Mongo DB. Mongo DB used as the knowledge base of the conversational bot where it stores the archeological sites information. Finally, the response passes to the user interface where user is capable to get the information required on archaeological sites in Sri Lanka.

D. Image Visualization using Augmented Reality

Fig. 5 shows the architectural overview of the image visualization component. In order to augment any image, first it is required to design three-dimensional model of the object. For that purpose, it was used Blender software [19]. Blender images are saved as ". Blend". It is able to provide different colors to different objects. Created license key for the image on which needs to augment the image. For that purpose, need an interface the image in Vuforia SDK [20]. Using development portal, able to create license key. Only the image of high-quality resolution can be augmented. High target image is created using Vuforia SDK and target manager is used to create image.

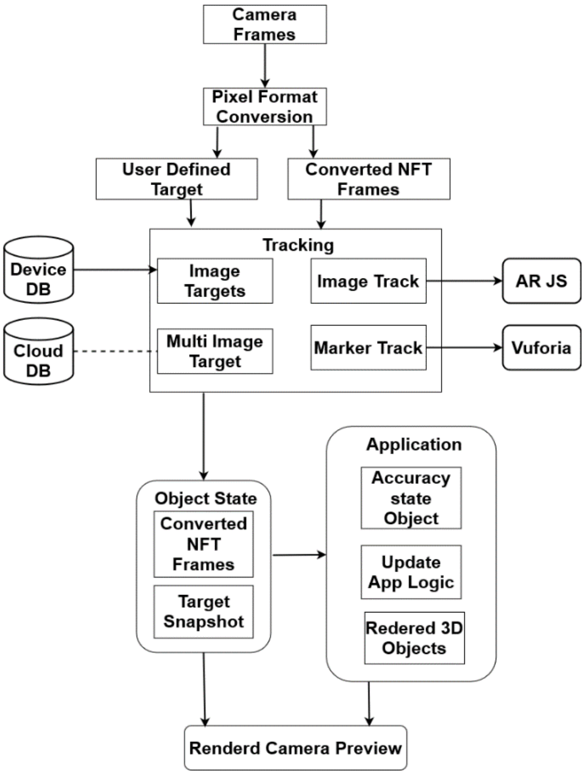

Fig. 5. Image visualization process overview

The 3D image created using Blender is imported in Unity software. The image would be inserted in AR camera. It could scale, rotate, and give position to the given image using Unity software. It could add 3D text in Unity 3D. Interface implement for any object; need to write C# script. Then, it is imported and interfaced using C# script. Android Studio is interfaced with Unity 3D software to build application. The build function creates ".apk" file, which can be easily downloaded to create an Android application. The mobile interfaces were created with the use of Unity3D.

The 3D model was integrated into the 3D interface as proposed. The augmented reality interface and the tracking procedure was created with the use of Vuforia architecture. The images of the existing archaeological sites in Sri Lanka were used as 3D image trackers and using Vuforia, it was made track able. Then the 3D model was placed within the field of view of the augmented reality camera inside of the interface. All the relevant functions were written in C# and embedded into Unity game objects.

IV. Experiment Results and Discussion

This section describes the experiments have been conducted to evaluate each module and the analysis of the given results.

A. Results Analysis of Image Identification

Highest accuracy level was given by the CNN model for the image identification module. The accuracy score is 0.98 as shown in the Fig. 6 below. Table I shows the experiment results conducted to select the best model.

TABLE I. EVALUATION OF IMAGE CLASSIFICATION MODELS

Deep Model	Accuracy
CNN	**0.94**
SVM	0.76

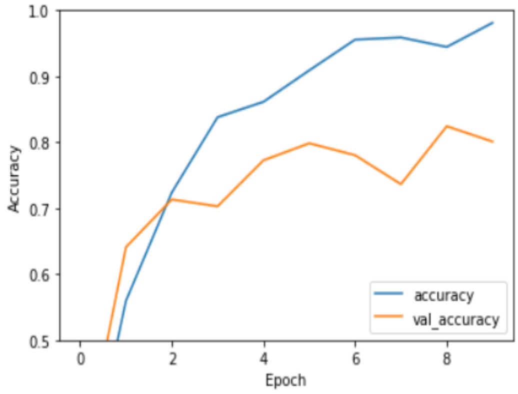

Fig. 6. Accuracy given by the image processing module

B. Results Analysis of the Chatbot

Chatbot is evaluated using following evaluation metrics. It used the correctly recognized intents and total number of intents to calculate the accuracy (Ac). Precision (Pr) denotes how many true positives over the summation of true positives and false positives. F1-score (F1) measures the accuracy on the model's classification tasks.

Dialog management module of the chatbot was evaluated over Keras policy and Embedding policy to select the most accurate policy configuration. Following Table II shows the experiment results given by the dialog management module for both of the policies on conversation level and action level.

TABLE II. DIALOG MANAGEMENT MODULE EVALUATION

Policy	Level	F1	Pr	Ac
Keras	**Conversation**	**0.973**	**0.962**	**0.947**
Embedding	Conversation	0.871	0.921	0.867
Keras	**Action**	**0.953**	**0.962**	**0.951**
Embedding	Action	0.913	0.873	0.882

C. Results Analysis of the Text Sumarization Module

TABLE III. ACCURACY RESULTS FOR ENCODER-DECODER ALGORITHM

LSTM Encoder-Decoder with Attention	**89.91**
LSTM Encoder-Decoder	78.69

Table III shows the accuracy results given for the Encoder-Decoder and Encoder-Decoder with Attention. It shows the highest accuracy was achieved by using Attention with Encoder-Decoder.

D. Results of the Augmented Reality Visualization

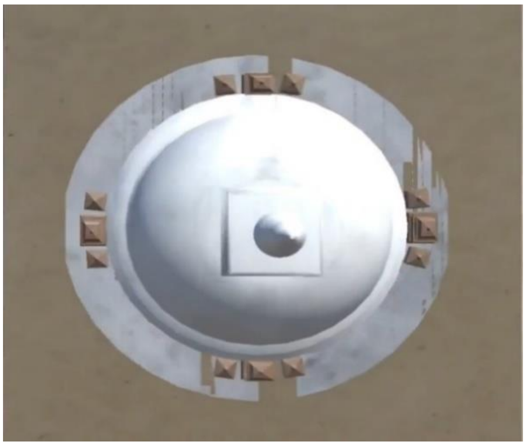

Fig. 7. 3D visualized view of Kiri Vehera

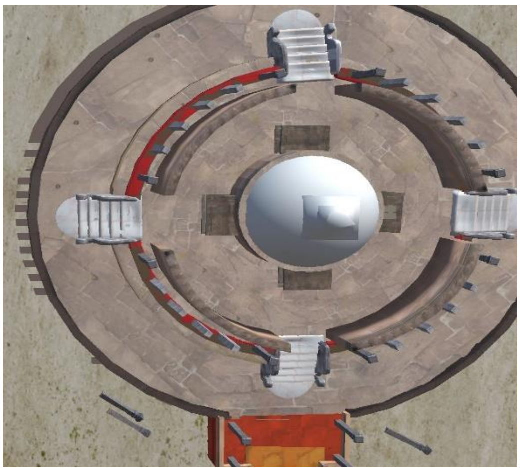

Fig. 8. 3D content creation for Vatadage

Fig. 7 shows the 3D visualization of the Kiri Vehera. Fig. 8 shows the 3D content creation for Vatadage.

CONCLUSION

The proposed system contributes to establish a smart platform for Sri Lankan Archaeological sites by using Artificial Intelligence & Machine Learning concepts. Initial concept was chosen Sigiriya, Vatadage, Hetadage, Kirivehera and Samadhi Statue as the places. It is assisted using four major components, namely: the image identification component, community platform and information base, artificial intelligence chatbot and the image visualization module. This platform helps to overcome several issues identified by the literature survey and referring the existing online resources in the archaeological domain. It provides an innovative solution for the archeology context in Sri Lanka. This platform would be further enhanced by integrating more smart features and choosing a wide range of archaeological sites.

REFERENCES

[1] A. Bandara *et al.*, "Lightning Risk Analyses on Protected Sri Lankan Heritage Stupas," *2018 13th Int. Conf. Ind. Inf. Syst. ICIIS 2018 - Proc.*, no. 978, pp. 322–327, 2018, doi: 10.1109/ICIINFS.2018.8721399.

[2] European Environment Agency (EEA), "Progress Report- 2018 Ministry of Tourism Development, Wildlife and Christian Religious Affairs," vol. 53, no. 9, pp. 1689–1699, 2019, doi: 10.1017/CBO9781107415324.004.

[3] "Sri Lanka - UNESCO World Heritage Centre." [Online]. Available: http://whc.unesco.org/en/statesparties/lk. [Accessed: 01-Sep-2020].

[4] "Archaeology & History Destinations, Tours and Exhibitions Around the World." [Online]. Available: https://archaeology-travel.com/. [Accessed: 06-Sep-2020].

[5] "Welcome to the official Website of the Department of Archaeology." [Online]. Available: http://archaeology.gov.lk/. [Accessed: 01-Sep-2020].

[6] B. Cui, W. Zhou, G. Fan, Y. Wu, and Y. Cui, "A smart and active way for protection of local cultural heritages," *2018 13th Int. Conf. Comput. Sci. Educ.*, no. Iccse, pp. 1–5, 2018.

[7] D. M. R. Kulasekara and S. M. B. Harshanath, "Detecting Changed Areas in Images from Different View Points," *Proc. 8th Int. Conf. Comput. Sci. Educ. ICCSE 2013*, pp. 525–528, 2013, doi: 10.1109/ICCSE.2013.6553966.

[8] A. Saini, T. Gupta, R. Kumar, A. K. Gupta, M. Panwar, and A. Mittal, "Image based Indian monument recognition using convoluted neural networks," *2017 Int. Conf. Big Data, IoT Data Sci. BID 2017*, vol. 2018-Janua, no. December, pp. 138–142, 2018, doi: 10.1109/BID.2017.8336587.

[9] R. Vijay, B. Vangara, and S. P. Vangara, "A Hybrid Model for Summarizing Text Documents Using Text Rank Algorithm and Term Frequency A Hybrid Model for Summarizing Text Documents Using Text Rank Algorithm and Term Frequency," no. June, 2020.

[10] E. U. Shankari, N. V. K. Rao, C. Muni, K. Sri, M. Spoorthi, and V. M. Datta, "Multi Document Text Summarization Using Genism Multi Document Text Summarization Using Genism," no. July, 2020.

[11] F. Clarizia, F. Colace, M. De Santo, M. Lombardi, F. Pascale, and D. Santaniello, "A Context-Aware Chatbot for Tourist Destinations," *Proc. - 15th Int. Conf. Signal Image Technol. Internet Based Syst. SISITS 2019*, pp. 348–354, 2019, doi: 10.1109/SITIS.2019.00063.

[12] M. Lombardi, F. Pascale, and D. Santaniello, "An application for Cultural Heritage using a Chatbot," *2nd Int. Conf. Comput. Appl. Inf. Secur. ICCAIS 2019*, pp. 1–5, 2019, doi: 10.1109/CAIS.2019.8769525.

[13] A. N. Jayawardena and I. Perera, "A framework for Mixed Reality application development: A case study on Yapahuwa archaeological site," *16th Int. Conf. Adv. ICT Emerg. Reg. ICTer 2016 - Conf. Proc.*, no. March 2018, pp. 186–192, 2017, doi: 10.1109/ICTER.2016.7829917.

[14] J. Keil, M. Zöllner, M. Becker, F. Wientapper, T. Engelke, and H. Wuest, "The House of Olbrich An Augmented Reality tour through architectural history," *2011 IEEE Int. Symp. Mix. Augment. Real. - Arts, Media, Humanit. ISMAR-AMH 2011*, pp. 15–18, 2011, doi: 10.1109/ISMAR-AMH.2011.6093651.

[15] G. D. D. M. Galmangoda, P. Y. S. Gajanayake, K. P. Indika, N. R. Rajapaksha, and Y. Jayaweera, "Augmented Reality to Reconstruct Sri Lankan Cultural Heritage in Prime State: HeladivaAR," *Symp. Stat. Comput. Model. with Appl.*, pp. 40–44, 2016.

[16] D. Greene and P. Cunningham, "Practical solutions to the problem of diagonal dominance in kernel document clustering," *ACM Int. Conf. Proceeding Ser.*, vol. 148, no. 2004, pp. 377–384, 2006, doi: 10.1145/1143844.1143892.

[17] C. Su, R. Gupta, S. Ananthakrishnan, and S. Matsoukas, "A Re-Ranker Scheme for Integrating Large Scale NLU Models," *2018 IEEE Spok. Lang. Technol. Work. SLT 2018 - Proc.*, pp. 670–676, 2019, doi: 10.1109/SLT.2018.8639519.

[18] T. Bocklisch, J. Faulkner, N. Pawlowski, and A. Nichol, "Rasa: Open Source Language Understanding and Dialogue Management," pp. 1–9, 2017.

[19] T. M. Takala, M. Makarainen, and P. Hamalainen, "Immersive 3D modeling with Blender and off-the-shelf hardware," *IEEE Symp. 3D User Interface 2013, 3DUI 2013 - Proc.*, pp. 191–192, 2013, doi: 10.1109/3DUI.2013.6550243.

[20] G. Tian-Han, T. Qiao-Yu, and Z. Shuo, "The Virtual Museum Based on HoloLens and Vuforia," *Proc. - 2018 4th Annu. Int. Conf. Netw. Inf. Syst. Comput. ICNISC 2018*, pp. 382–386, 2018, doi: 10.1109/ICNISC.2018.00084.