Informatics Bulletin, Faculty of Computers and Artificial Intelligence, Helwan University
Published Online ***Vol 5 Issue 2, July 2023***
(https://fcihib.journals.ekb.eg)

Survey Study: Monument Recognition using Artificial Intelligence

Mennat Allah Hassan[1,2,*], Alaa Hamdy[1], Mona Nasr[2]

[1]*Faculty of Computer Science, Misr International University, Cairo, Egypt.*
[2]*Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt.*
mennatallah.sayed@miuegypt.edu.eg, *alaa.hamdy@miuegypt.edu.eg*, *m.nasr@helwan.edu.eg*

(a) (b) (c) (d)

Fig. 1. Examples of monuments in different art-forms from: (a) India [7], (b) France [27], (c) Greece [31], (d) Egypt [8].

***Abstract*— Monuments are man-made structures that are designed to imbue a location with important meanings. They serve both a political and artistic function. There is an overwhelming amount of information and features to look into monuments. Monument work is a unique industry. It would help the economies and growth of many countries. Every culture generally has its own set of distinct characteristics, like monuments, writing, and music. Considering the history and stories associated with monuments, it is difficult to evaluate the techniques used to detect these monuments. Within the domain of object recognition and classification, monument recognition is an arduous task. Numerous difficulties must be overcome, as multiple variables can alter the recognition method. This paper has been written to study the recent research on this point from 2017 to 2023. In this study, monuments from a variety of civilizations and countries, including Egypt, India, Turkey, Paris, Greece, and Singapore, have been compared to the algorithms or methodology utilized.**

***Index Terms*—computer vision, deep learning, landmark, machine learning, monument recognition**

I. Introduction

One of the largest industries in the world is tourism. With 7% of global trade in 2019, it is the third-largest export category after fuels and chemicals. It can represent more than 20% of the gross domestic product in some countries; It is the third-most exported industry globally [1]. One of Egypt's most significant sources of income is the tourist sector. It brings in a lot of money annually in both dollars and foreign currencies, making it a big part of the country's gross domestic product. Egypt is widely recognized as one of the most popular tourist destinations across the world, with a significant number of visitors traveling to the country annually. It comes out because it has so various types of tourist attractions, like temples, museums, monuments, historical buildings, and buildings with artistic or historical value. Egypt has one-third of all the monuments in the world [2].

The UNESCO committee's announcement in 2021 regarding the top countries of World Heritage locations highlights France as the fifth-ranked country with 49 world heritage locations and India as the sixth-ranked country with 40 world heritage locations. Additionally, Turkey and Greece are noted with Turkey boasting 19 and Greece boasting 18 World Heritage locations [3].

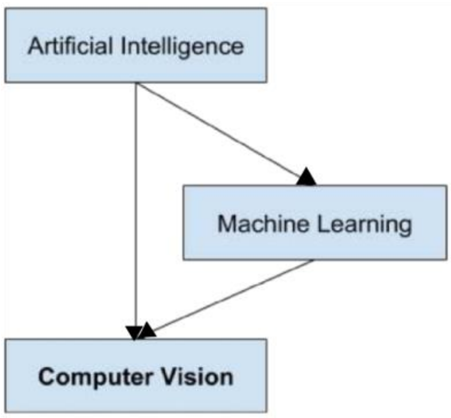


Fig. 2. Overview of the relationship of AI and CV [9].

Monuments are constructed forms that are built to give prominent space meanings. Monuments have both an artistic value and a political purpose. Typically, political elites create monuments to promote selected historical realities that highlight convenient events and people [4]. So, a monument is a statue, building, or other structure created to honor a remarkable individual or event [5]. Bridge, Bust, Cross, Equestrian Monument, Fountain, Mausoleum, Obelisk, Pyramid, Reliquary, Sarcophagus, Statue, Stele, Tomb, and Triumphal arch [6] are examples of monument types, as seen in Fig. 1. Naturally, people can recognize objects in general and monuments, in particular, using their human eyes and brains. They learn from books or other resources that tell them what the monument is and the story behind each unique one; This is called self-learning. For instant recognition, they may obtain information from experts such as monument scientists or tour guides. Currently, humans can recognize monuments using computer vision (CV).

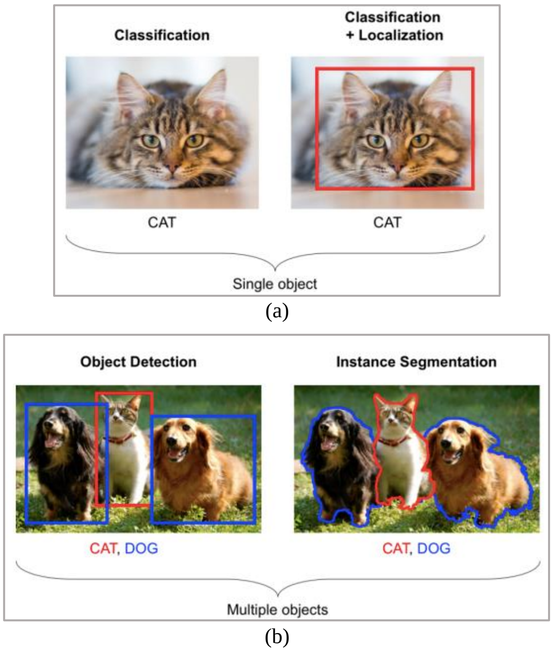


Fig. 3. A Closer look of detection meaning [10]. (a) Single object in image. (b) Multiple objects in image.

CV is a field of study concerned with assisting computers to see. The field of CV is interdisciplinary that might be considered a subcategory of artificial intelligence (AI) and machine learning (ML); It includes specialized methodologies and general learning algorithms, as seen in Fig. 2. It seeks to comprehend the content of digital images. Usually, this requires the development of techniques that seek to simulate human vision. It may be essential to extract a description from a digital image in order to understand its content. This description could be of an object, a textual explanation, a three-dimensional model, or something else. Numerous well-known computer vision applications seek to recognize things in digital photos and videos; For example, object detection concerns where the things are in the photograph. Object Landmark Detection focuses on identifying the photographed object's focal points. Object recognition is what and where the objects are depicted in this photograph [8]. In general, detection or localization is a task that identifies an object in an image and localizes it using a bounding box. This task has numerous applications, including locating pedestrians and signboards for autonomous vehicles. Fig. 3 below illustrates detection [9].

Various CV applications, such as monitoring, robotics, and human interaction, often include object detection. Nevertheless, object recognition is extremely tough and challenging due to complicated background, illumination variation, scale variation, occlusion, and object deformation [10]. The effectiveness of objection detection seems stable in 2010. Despite the fact that various solutions are still being presented, performance enhancement is quite limited. Currently, deep learning (DL) is beginning to outperform traditional computer vision techniques in certain areas of computer vision. With the tremendous success of DL in image categorization [13, 14, 15, 16], researchers are beginning to investigate how DL may be used to improve object recognition performance. Object detection based on DL has also made significant progress over the past few years [17, 18, 19].

In CV, the promise of DL is improved performance by models that may require more data but less knowledge of digital signal processing to train and operate. Beyond the hype and grand claims around deep learning methods, these methods attain state-of-the-art performance on complex challenges. That challenges are particularly in computer vision tasks like image classification, object recognition, and face detection. Some of the earliest large-scale demonstrations of deep learning's power were in computer vision, notably image recognition. Presently, they have been in object detection and face recognition. The following is a list of the five promises that can be realized by the application of DL technologies in CV [13]:

• The Potential for Fully Automated Feature Extraction; Raw image data can be utilized for automatic learning and feature extraction.

• The Potential of End-to-End Models; Single models that cover the entire process can take the role of multiple specialized model pipelines.

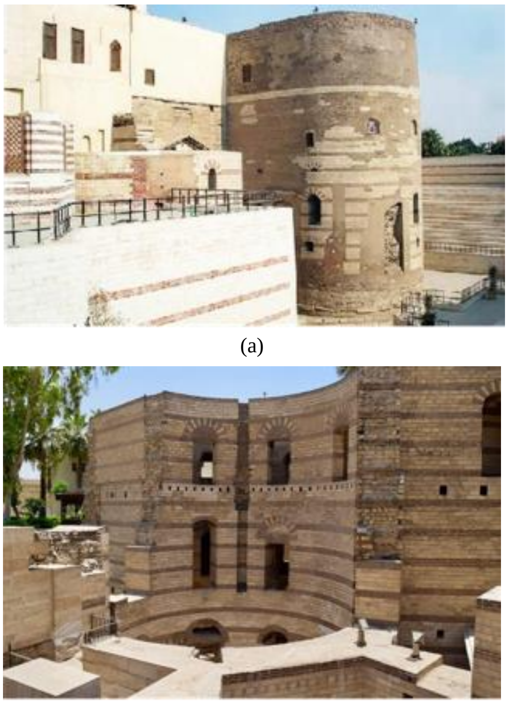
• The Potential of Multiple Model Uses; It is possible to reuse learned features and even complete models in different contexts.

• The Potential of Outstanding Performance; Applying techniques demonstrates a higher level of ability when dealing with difficult jobs.

• The Prospects Offered by the General Method; A variety of related jobs can all be tackled with the help of a solitary overarching method.

Deep convolutional neural networks (CNN) are the foundation of the most cutting-edge techniques for deep object detection that are available today [17, 20, 21, 19].

CNNs, which are members of the Neural Network family, are Deep Learning techniques that accept images as input. It assigns emphasis (weights) to particular properties of an image's objects and attempts to discern between them [22]. Considering the last trait, it can be deduced that CNN could likewise provide improved results for monument recognition [23].



(a)

(b)

Fig. 5. Example of different view's angles for Babylon Fortress [13]. (a) Babylon Fortress as a front view. (b) Babylon Fortress as a back view.

Monument recognition is unquestionably an important element for both tourists and locals in many parts of the world. Both can increase their knowledge of their respective cultural traditions. Consequently, a person can download an application to their smartphone and then activate the monument identification application. Hence, he acquires information about the monument in front of them [24] even if no guides are available.

Image-based monument recognition presents a variety of issues for monument recognition. The complexity which gives a precise definition of what is and what is not a monument; This is another challenge, as explained in Fig. 4. Furthermore, the challenge of the differences in the photos of a particular monument is due to their orientation, as shown in Fig. 5. Moreover, image transformations— scaling, translation, and rotation— when capturing the images are an important problem in Fig. 6. In addition, the occlusion problem, which is prevalent in photos of trees, people, animals, and other objects, reduces accuracy, as illustrated in Fig. 7.

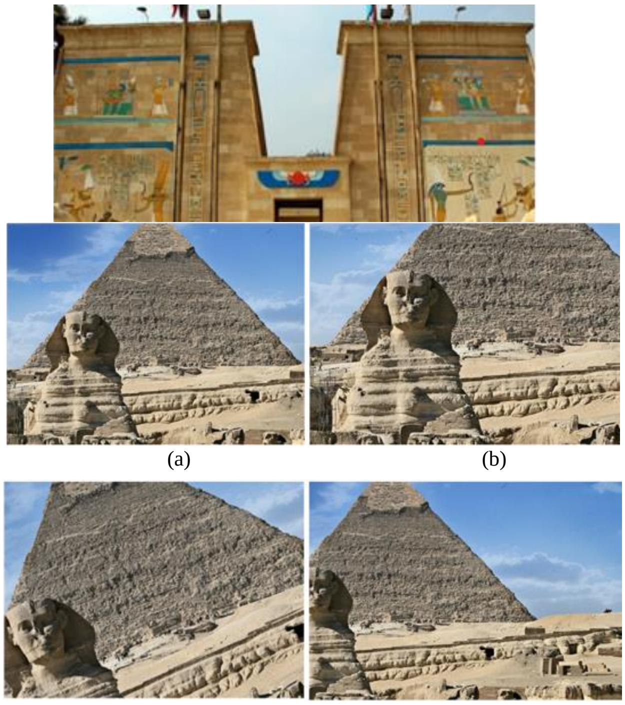
II. Literature Survey

Through literature review, many studies were done on applying intelligent techniques such as Residual Network (ResNet50), MobileNet, AlexNet, DCNN, and more for monument detection. Different datasets have been used in their studies, as shown in Table 1.

M. N. Razali et al. [36] proposed a robust and lightweight landmark recognition model in 2023. This study employed a hybrid model that incorporates CNN and Linear Discriminant Analysis techniques. The research evaluated the effectiveness of this approach on two distinct datasets, namely the Universiti Malaysia Sabah (UMS) landmark and the Scene-15 datasets. The outcomes demonstrated that the most effective feature extraction and classification method was the EFFNET-CNN. Results of this approach yielded 100% and 94.26% accuracy in classification for the UMS landmark and Scene-15 datasets, separately.

K. Yasser et al. [28] introduced a system for recognizing monuments and provided a thorough description in 2022. This research addressed the difficulty of recognizing cropped monuments as a significant challenge. Their primary contribution was applying generative adversarial deep learning algorithms (GAN) to outpaint the cropped image. The outpainted image was then sent into the state-of-the-art classifier RESNET, which classified the monument and displayed a thorough description of its extraordinary history. The system was trained utilizing data collected by the authors' team. After 1000 training epochs, the training GAN's Adversarial Loss is 0.28344182, and its validation loss is 0.30181705. During testing, the RESNET classifier achieved an accuracy rate of 97.0%, a precision rate of 97.1%, a recall rate of 97.0%, and an F1-measure of 97.0%. Their limitations were inpainting in images. In addition, the improvement of the classification procedure.

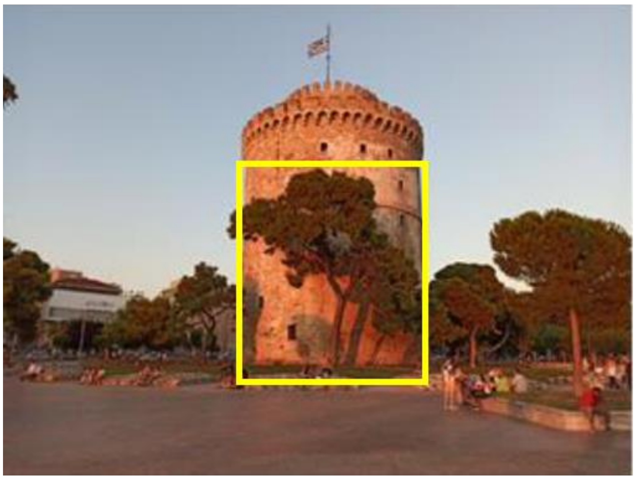
In [29] that same year, M. Trivedi emphasized that the postures of 3D object monuments are problematic. This issue is



(a) (b)

(c) (d)

Fig. 6. Example of different transformations for Pyramid of Khafre and Sphinx [14]. (a) Original image. (b) Scaled-in image. (c) Rotated image. (d) Translated image— during capturing the images.



(a)

(b)

Fig. 7. Sample of occlusion problem [31], (a) The White Tower in Greece as a far version. (b) The White Tower in Greece as a close version.

a result of divergent perceptions of the images of the monuments. Depending on the angle from which the photograph was taken, there could be a big variation. In this work, data augmentation is performed to address this issue. Flip, Rotation, Translation, Zoom, Contrast, Hue, Brightness, and Saturation were applied to the dataset as part of the data augmentation process. This work deployed DL architectures on a hybrid dataset comprised of the "Indian Monument Recognition Dataset" and "Qutub Complex Monuments' Images Dataset". In recognition of the monuments, the accuracy of InceptionV3, MobileNet, ResNet50, AlexNet, and VGG16 was 97.79%, 93.73%, 86.47%, 68.88%, and 61.33%, respectively.

S. Hesham et al. [30] applied three algorithms to compare ML and DL in the context of monument recognition in 2021 on an Indian dataset from Kaggle that includes 1,286 images. These algorithms were ResNet50, VGG16, and KNN. Resnet performed better than VGG16 and KNN at classifying not seen data, as measured by performance metrics. Their constraint was the limited number of classes of monuments, only three.

During the year 2021, T. Stougiannis [31] diligently gathered a grand total of 4,708 images featuring 18 diverse UNESCO World Heritage Monuments. They may be found in the city of Thessaloniki. The author used Tensorflow's Object Detection API to train a CNN detection model and then used Tensorflow Lite to integrate the model into an Android application. In addition to this, he contributed to the development of the field of data augmentation as a technique for improving the effectiveness of his model.

B. Bayram [32] compared in the earlier year the VGG, ResNet, and DenseNet DL architectures for the recognition of ten Istanbul, Turkey, historical landmarks. The Maiden's Tower, the Sultan Ahmet Mosque (Blue Mosque), the Galata Tower, Hagia Sophia, the Ortak oy Mosque, the Topkapi Palace, the Valens Aqueduct, the Dolmabahce Palace, the Obelisk of Theodosius, and the Dolmabahce Clock Tower are the 10 historical landmarks. The experimental findings demonstrated that the DenseNet-169 architecture was highly effective for the given dataset. It achieved an accuracy of 96.3%.

TABLE I
COMPARATIVE RESULTS

Reference, Year	Dataset	Approach/ Methodology	Results
[36], 2023	Two datasets: - UMS landmark dataset - Scene-15 dataset	EFFNET-CNN	100%, 94.26%
[28], 2022	Egyptian monuments Manually collected,	GAN (Prepossessing) ResNet50 (Classification)	GAN's Adversarial Loss= 0.28344182, its validation loss is 0.30181705. 97.0%
[29], 2022	Hybrid dataset	InceptionV3, MobileNet, ResNet50, AlexNet, VGG16.	97.79%, 93.73%, 86.47%, 68.88%, 61.33%.
[30], 2021	Indian monuments, Kaggle website	ResNet50 VGG16 KNN	88%. 83%. 64%.
[31], 2021	Manually collected, Thessaloniki City, Greece.	SSD MobileNet V1 coco, SSD MobileNet V2 coco.	93.46% 95.66%.
[32], 2020	Collected images of Istanbul, Turkey. Two different datasets: -Istanbul-2500 -Istanbul-5000	VGG16 VGG19 ResNet50 ResNet101 ResNet152 DenseNet121 DenseNet169 DenseNet201	(Istanbul-5000) 83.1% 83.6% 93.2% 92.4% 81.4% 96.1% 96.3% 94.4%
[26], 2019	46 different monuments	MobileNet	95%.
[33], 2019	Paris	CNN	sensitivity— images with landmarks, 80% without geo info - 92% with geo info. specificity — images without landmarks, 99% without geo info - 99.5% with geo info.
[27], 2018	Two datasets: - Singapore landmarks - Paris landmarks	AlexNet, GoogLeNet, SqueezeNet, NU-LiteNet-A, NU-LiteNet-B.	(Singapore - Paris) 64.82% - 58.62% 70.69% - 59.97% 60.08% - 53.34% 78.09% - 66.67%

			81.15% - 69.58%
[34], 2018	Dataset includes Indian Mughal Monuments Great Cathedrals.	CNN	80%
[35], 2017	Manually collected, Indian monuments	DCNN-using fc6 layer.	92.7%
[25], 2017	Used web-crawler, Indian Monuments	Inception v3 architecture.	96-99%

The author of V. Palma [26] presented a study investigating CNN techniques in the context of architectural heritage. A research area that is in the process of being developed. The relationship between DL algorithms and modern information modelling was studied to promote legacy collections and new object identification methodologies. The algorithm implemented by CNN is based on the MobileNet model. The dataset comprised between 50 and 100 images for each of the 46 monuments. They utilized data augmentation techniques to enrich the 500-image per monument dataset. The model was trained to recognize the desired landmark with at least 95% accuracy. On the basis of a test subset of photos, it was calculated that the total accuracy of the trained models exceeded 95%.

A. Boiarov and E. Tyantov [33] presented an innovative approach for landmark recognition in images, which they had previously tested and successfully implemented at Mail.ru. The application of this technology allowed the identification of well-known places, buildings, monuments, and other types of landmarks within user photographs. The fact that it was exceedingly difficult to give a specific definition of what constitutes a landmark and what does not include a landmark was the primary obstacle that they faced. Certain buildings, monuments, and natural features are considered landmarks, while others are not. They utilized two metrics to measure the outcomes of the experiments: Sensitivity is defined as the accuracy of a model on images that contain landmarks, and it is defined as having a value of 80%. Specificity is defined as the accuracy of a model on images that do not contain any landmarks, and it is defined as having a value of 99%.

In their paper [27], C. Termritthikun and colleagues discussed the creation of NU-LiteNet, a new CNN model, and its design. They created the NU-LiteNet CNN model based on the SqueezeNet development concept. Three models required the least amount of processing time: NU-LiteNet-A (637 ms), NU-LiteNet-B (706 ms), and SqueezeNet (773 ms). These models have the smallest model sizes: NU-LiteNet-A of 1.07 MB, SqueezeNet of 2.86 MB, and NU-LiteNet-B of 2.92 MB. As a result, NU-LiteNet-A was the most time- and space-efficient model.

Cathedrals and Indian Mughal Monuments were the two categories used to categorize the images in the proposed work by A. Ninawe et al. [34]. They used CNN, which is a software library that is open-source and is based on TensorFlow. The significant endeavor that needed to be accomplished, was to extend the model to work effectively on the new dataset. Caffe, a framework used for deep learning, was utilized so that trained CNN weights could be obtained. They trained CNN with a dataset of 5,000 images from Indian and Mughal structures, cathedrals, and churches. Python 2.7 and the tensor flow, especially Keras, are used in the computation of every image for training and testing. They came up with an innovative and potentially fruitful model for image classification that had an accuracy of higher than 80. The used Python version is concerned with a drawback; Python 2.7 of TensorFlow is not supported anymore.

A. Saini et al. [35] presented a method for classifying numerous monuments that were derived from the characteristics of picture representations of the monuments. In order to extract representations, cutting-edge Deep Convolutional Neural Networks (DCNN) were utilized. The dataset was obtained through the manual collection and consisted of one hundred folders, each of which had fifty images of a different landmark. The DCNN approach resulted in an accuracy rate of 92.7% across the board.

S. Gada et al. [25] utilized a well-known Deep Learning architecture model to identify the images in a timely manner while achieving remarkable levels of accuracy. They used a previously retrained model and collected around 400 pictures of each landmark. After 4000 iterations, they successfully achieved a cross-entropy that was near 0.067, an accuracy in training of 99.4%, and an accuracy in testing that was equivalent to training accuracy and predicted to be between 96% and 99%.

III. CONCLUSION

Monument recognition is an expanding field in the digital age. In this study, we conduct a literature review on monument recognition. As seen in Table 1, each publication utilized a unique dataset and a distinct set of methodologies. Here we are continuing to address the difficulties of Monument recognition.

Based on previous research, it is evident that there is still a significant opportunity to train advanced deep-learning models. These models would be capable of detecting numerous visual objects regardless of their attributes or viewpoints. In terms of performance and processing capability, there is still the potential for employing DL models to detect things on smartphones in real-time.

REFERENCES

[1] UNWTO, “Tourism and COVID-19—Unprecedented Economic Impacts,” 2020. [Online]. Available: https://www.unwto.org/tourism-and-covid-19-unprecedented-economic-impacts, Accessed on: Feb. 27, 2022.

[2] A. Derbali, “The importance of tourism contributions in Egyptian economy,” International Journal of Hospitality and Tourism Studies (IJHTS), Jun. 2020.

[3] “World Heritage List Statistics,” UNESCO World Heritage Centre, 2021. [Online]. Available: https://whc.unesco.org/en/list/stat. [Accessed: Mar-2022].

[4] F. Bellentani, and M. Panico, “The meanings of monuments and memorials: toward a semiotic approach. Punctum,” International journal of semiotics, 2(1), 28-46, 2016.

[5] Oxford University Press. (n.d.) Emotional Intelligence. “Oxford English dictionary,” [Online]. Available: https://www.oed.com/ Accessed on: Jan. 15, 2022.

[6] M. Stocker, “Public monument,” 2003.

[7] V. Sharma, “Qutub Complex Monuments' Images Dataset,” Kaggle, Oct-2018. [Online]. Available:

https://www.kaggle.com/datasets/varunsharmaml/qutub-complex-monuments-images-dataset. [Accessed: Mar-2022].

[8] M. A. Hassan, "Egypt Monuments Dataset v1," GitHub. [Online]. Available: https://github.com/mennatallahhassan/egypt-monuments-dataset. [Accessed: Mar-2023].

[9] J. Brownlee, "Deep learning for computer vision: image classification, object detection, and face recognition in python," Machine Learning Mastery. 2019.

[10] R. Shanmugamani, "Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras," Packt Publishing Ltd, 2018.

[11] "Pharaonic Village," Egypt State Information Service (SIS), 24-Apr-2016. [Online]. Available: https://www.sis.gov.eg/Story/101114/Pharaonic-Village?lang=en-us. [Accessed: 24-Sep-2021].

[12] E. Cummins, "Temple of Amun-Re and the Hypostyle Hall, Karnak," Khan Academy. [Online]. Available: https://www.khanacademy.org/humanities/ap-art-history/ancient-mediterranean-ap/ancient-egypt-ap/a/karnak. [Accessed: 24-Sep-2021].

[13] D. Yasser and S. Refaat, "Egyptian Monuments, Religion Complex in Old Egypt," Kaggle, Mar-2021. [Online]. Available: https://www.kaggle.com/datasets/daliayasser/egyptian-monuments. [Accessed: 27-Feb-2022].

[14] A. Kooser, "How did Egyptians move pyramid stones? Mystery may be solved," CNET. [Online]. Available: https://www.cnet.com/science/physicists-figure-how-ancient-egyptians-moved-pyramid-stones/. [Accessed: 09-Sep-2021].

[15] X. Jiang, A. Hadid, Y. Pang, E. Granger, and X. Feng, "Deep learning in object detection and recognition," Springer Singapore, 2020.

[16] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L.Fei-Fei, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee, Jun. 2009.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778), 2016.

[18] K. Simonyan, , and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[19] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, ... and A. Rabinovich, "Going deeper with convolutions," in proceedings of the IEEE computer society conference on computer vision and pattern recognition, 2015.

[20] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587), 2014.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition,". EEE transactions on pattern analysis and machine intelligence, 37(9), 1904-1916, 2015.

[22] R. Girshick, and R. Faster, "C. N. N," Towards real-time object detection with region proposal networks," arXiv preprint arXiv:1506.01497, 2015.

[23] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in Proceedings of the IEEE international conference on computer vision (pp. 2961-2969), 2017.

[24] T. Y. Lin, P. Goyal, R. Girshick, K. He, & P. Dollár, "Focal loss for dense object detection," in Proceedings of the IEEE international conference on computer vision (pp. 2980-2988), 2017.

[25] S. Gada, V. Mehta, K. Kanchan, C. Jain, and P. Raut, "Monument recognition using deep neural networks," in 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) (pp. 1-6). IEEE, Dec. 2017.

[26] V. Palma, "TOWARDS DEEP LEARNING FOR ARCHITECTURE: A MONUMENT RECOGNITION MOBILE APP," International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 2019.

[27] C. Termritthikun, S. Kanprachar, and P. Muneesawang, "NU-LiteNet: Mobile landmark recognition using convolutional neural networks," arXiv preprint arXiv:1810.01074, 2018.

[28] K. Yasser, A. M. Salama, A. Amr, L. E. Yehia, S. Refaat, and F. H. Ismail, "Egyart_classify: an approach to classify outpainted Egyptian monuments images using GAN and ResNet," in 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) (pp. 160-166). IEEE, May 2022.

[29] M. Trivedi, S. Agrawal, A. Kumar, S. K. Thakur, and N. Gautam, "Indian Monument Recognition Using Deep Learning," in ECS Transactions, 107(1), 15563, 2022.

[30] S. Hesham, R. Khaled, D. Yasser, S. Refaat, N. Shorim, and F. H. Ismail, "Monuments recognition using deep learning vs machine learning," in 2021 IEEE 11th annual computing and communication workshop and conference (CCWC) (pp. 0258-0263). IEEE, Jan. 2021.

[31] T. Stougiannis, "Landmark and monument recognition with Deep Learning," M.S. thesis, School Of Science & Technology, International Hellenic Univ., Thessaloniki, Greece, 2021.

[32] B. Bayram, B. Kilic, F. ÖZOĞLU, F. ERDEM, T. Bakirman, S. Sivri, ... and A. Delen, "A Deep learning integrated mobile application for historic landmark recognition: A case study of Istanbul," Mersin Photogrammetry Journal, 2(2), 38-50, 2020.

[33] A. Boiarov, and E. Tyantov, "Large-scale landmark recognition via deep metric learning," in Proceedings of the 28th ACM International Conference on Information and Knowledge Management (pp. 169-178), Nov. 2019.

[34] A. Ninawe, A. K. Mallick, V. Yadav, H. Ahmad, D. K. Sah, and C. Barna, "Cathedral and Indian Mughal Monument Recognition Using Tensorflow," in International Workshop Soft Computing Applications (pp. 186-196). Springer, Cham, Sep. 2018.

[35] A. Saini, T. Gupta, R. Kumar, A. K. Gupta, M. Panwar, and A. Mittal, "Image based Indian monument recognition using convoluted neural networks," in 2017 International Conference on Big Data, IoT and Data Science (BID) (pp. 138-142). IEEE, Dec. 2017.

[36] M. N. Razali, E. O. N. Tony, A. A. A. Ibrahim, R. Hanapi, and Z. Iswandono, "Landmark Recognition Model for Smart Tourism using Lightweight Deep Learning and Linear Discriminant Analysis," in International Journal of Advanced Computer Science and Applications, 14(2), 2023.