

# Building and evaluation of an Algerian Cultural Heritage dataset using convolutional neural networks

1<sup>st</sup> Toufik Djelliout

Department of Computer Science, Faculty of Exact Sciences,  
University of Bejaia, Algeria  
Research Center for Scientific and Technical Information  
Algiers, Algeria

2<sup>nd</sup> Hassina Aliane

Research Center for Scientific and Technical Information  
Algiers, Algeria

**Abstract**—Preservation of cultural heritage is a field of high importance. Recently people are sharing architectural monument images on social media. In this paper, we try to recognize architectural monuments in digital photographs of Algerian cultural heritage using a convolutional neural network (CNN). As no datasets support the diversity of Algerian monuments, adapted for CNN training, we built a new dataset and made available to the public. AlgHeritage dataset consists of 20,000 images that can serve as a benchmark for various research fields, as it combines numerous real-world challenges. We evaluate our dataset with three CNN models MobileNetV3, InceptionV3 and InceptionResNetV2, and compare it with two other datasets. MobileNetV3 with fine-tuning produces a good accuracy of 93.29% on the AlgHeritage dataset in comparison with other datasets. The AlgHeritage dataset is available at <https://bit.ly/3O38FOe>.

**Index Terms**—Architectural monuments, cultural heritage, dataset, image recognition, convolutional neural network

## I. INTRODUCTION

For years, the preservation of tangible cultural heritage has been an important issue with social-cultural impact and technological development in the area of computer systems throughout the world.

In Algeria, the interest in digital cultural heritage has increased considerably since it is the repository of several civilizations, especially in architectural heritage. This heritage has particularly a rich past because it is diverse and cosmopolitan: Prehistoric, Phoenician, Berber, Roman, Vandal, Byzantine, Arabic, Ottoman, Spanish and French. Some Algerian architectural heritage sites are in danger due to poor management, extensive new urban development, and general neglect. For example, the Algiers Casbah which was added to the list of World Heritage by UNESCO [1] is at risk of damage [2]. For that, experts have resorted to digital tools to attempt the conservation and preservation of endangered and damaged sites. Thousands of pictures are being uploaded due to the emergence of social networks such as Facebook and Twitter. Nevertheless, there are no datasets for the Algerian monuments heritage, including the need to preserve this heritage.

We note that architectural style and period differences of monuments contribute to a high intra-class variance. For ex-

ample, Roman and Byzantine architectures are quite different but use the same monument types as arches or temples. Architectural monuments recognition is hence a challenging task which often struggles with the recognition of small details, the classification of images aims to resolve this issue.

In this article, we identify images of Algerian monuments heritage with deep learning (DL) techniques. The main objective is to experiment the use of the most representative CNN to identify the monuments heritage. This paper is organized as follows. Section 2 illustrates the literature related to datasets in architectural heritage. Section 3 shows the characteristics of the Algerian architectural structures in the AlgHeritage dataset and the details of the datasets which are used in the experiments. Section 4 presents the CNN models which provide our results in image recognition. In section 5, we present the experiment. Afterwards we mentioned the details of the results in section 6 and we finished with the conclusion in the last section.

## II. RELATED WORK

Many datasets have been used to measure the accuracy of image classification and image recognition. The dataset can serve as a reference for different research fields like medicine [3], [4], biology [5], plants [6], [7], foods [8], fish [9] and others; in our paper, we are interested in the recognition of the architectural cultural heritage monuments. For research, there are few datasets available in cultural heritage; each dataset is created and organized according to specific needs: architectural periods, architectural styles or architectural monuments.

Obeso created the MexCulture dataset for the first time to classify Mexican buildings into three classes according to the architectural periods [10]: prehispanic, colonial, and modern but this dataset is not available. After that, he created MexCulture142 dataset that is available [11] which is a subset of the MexCulture dataset, to identify 142 Mexican buildings, each building containing just 2 images. He uses a psycho-visual experiment to record gaze fixations of subjects executing a visual task of recognizing architectural styles of Mexican Cultural heritage. However, this dataset does not allow good