Proceedings of the joint international event 9th ARQUEOLÓGICA 2.0 & 3rd GEORES, Valencia (Spain). 26–28 April 2021

Received: 27/11/2020
Accepted: 26/03/2021

AUTOMATIC MASKS GENERATION FOR THE DIGITIZATION OF A WOODEN COLLECTION OF MAQUETTES USING DEEP LEARNING[1]

Giacomo Patrucco[a,*], Francesco Setragno[b]

[a] Laboratory of Geomatics for Cultural Heritage (G4CH Lab), Department of Architecture and Design (DAD), Politecnico di Torino, Viale Mattioli 39, 10125 Torino, Italy. giacomo.patrucco@polito.it

[b] Volta® A.I., Via Roberto Lepetit 34, 21040 Gerenzano VA, Italy. francesco@volta.ai

Abstract:

In the last few years, a new challenge for the researchers operating in the framework of 3D metric survey and modelling has been represented by the possibility to develop new strategies to digitize the artwork objects and the historical artefacts stored in our museums. The achievable 3D models can be used for many applications and they can effectively contribute to the management of the collections, to the dissemination and the sharing projects based on online 3D viewers. This appears evident in the historical period we are experiencing: due to the global pandemic limiting access to the places of culture, digital technologies can provide some valid alternatives to access heritage, improving the accessibility of these movable assets. Among the many technologies and techniques for the digitization of heritage objects, digital photogrammetry (usually based on Structure-from-Motion algorithms) can represent an efficient, effective, and relatively low-cost solution, since it allows to achieve high resolution 3D models with photographic textures characterised by high radiometric resolution. However, many procedures are still very repetitive and time-spending, requiring many resources from the operator in order to be fulfilled. Nowadays, deep learning can be a valuable instrument to improve the efficiency of these operations with a high degree of accuracy. The acquisition of the data used during the presented research has been carried out in the framework of the B.A.C.K. TO T.H.E. F.U.T.U.RE. project started in 2017, during which a collection of wooden maquettes representing Egyptian architectures has been digitized. For each of the 26 pieces, a complete photogrammetric survey has been performed using a fixed camera and a rotating platform. In order to process the images and to generate a texturized 3D model, an exclusion mask has been generated for each image (to exclude the background). This procedure was very time-consuming and, one of the main aims of this research was to establish an alternative workflow characterized by higher automatism.

A dataset of 3000 images (and corresponding exclusion masks) has been used to train the Neural Network. The images have been downsized due to their very high resolution (8688 x 5792 pixels) and augmentation strategies (random rotation; random cropping; horizontal and vertical flipping; random brightness change; random occlusions) have been applied, in order to increase the diversity of the dataset. 90% of the images have been used as a training set, while the remaining 10% has been used as a validation set (to verify the performance of the model during the training). After the training of the model, a post-processing phase has been designed to fix some topological errors produced by the segmentation process. Both 1-class scenario and 2-class scenario have been considered during this experience (using properly labelled dataset): in the first case, the Neural Network has been trained to recognize only the object of interest (the maquette), while in the second case it has been trained to perform a semantic segmentation and to classify the maquette as Class 1 and the metric bars (placed on the acquisition stage in order to provide metric reference during the photogrammetric process) as Class 2. As regards both scenarios Accuracy, IoU (Intersection over Union) and F1 score have been considered and evaluated. Starting from the segmented images, it was also possible to perform the 2-Class segmentation also on the 3D point cloud (each point has been labelled with the value of the corresponding Class) and the final result was characterised by an adequate level of accuracy. The obtained results have been presented and discussed (both quantitatively and qualitatively). In addition, in the conclusions, the next steps of the current research have been discussed and they will focus on improving the generalization capability of the model and on the possibility to use texturized 3D models to create an artificial dataset as an augmentation strategy for Neural Network training.

Keywords: close-range photogrammetry, deep learning, semantic segmentation, automatic masking, movable heritage

[1] Please check its full version at the *Virtual Archaeology Review*: MULTICLASS SEMANTIC SEGMENTATION FOR DIGITISATION OF MOVABLE HERITAGE USING DEEP LEARNING TECHNIQUES. https://polipapers.upv.es/index.php/var/article/view/15329

* Corresponding Author: Giacomo Patrucco, giacomo.patrucco@polito.it