HIGH RESOLUTION IMAGE PROCESSING AND LAND COVER CLASSIFICATION FOR HYDRO-GEOMORPHOLOGICAL HIGH-RISK AREA MONITORING

G. Miniello[1,3,a], M. La Salandra[2]

[1] *Department of Physics, University of Bari, Italy*

[2] *Department of Earth and Geoenvironmental Sciences, University of Bari, Italy*

[3] *Istituto Nazionale di Fisica Nucleare – Sezione di Bari, Italy*

E-mail: [a] giorgia.miniello@ba.infn.it

High-resolution image processing for land surface monitoring is fundamental to analyze the impact of different geomorphological processes on Earth surface for different climate change scenarios. In this context, photogrammetry is one of the most reliable techniques to generate high-resolution topographic data, being key to territorial mapping and change detection analysis of landforms in hydro-geomorphological high-risk areas. An important issue arises as soon as the main goal is to conduct analyses over extended areas of the Earth surface (such as fluvial systems) in a short time, since the need to capture large datasets to develop detailed topographic models may limit the photogrammetric process, due to the high demand of high-performance hardware. In order to investigate the best set up of computing resources for these very peculiar tasks, a study of the performance of a photogrammetric workflow based on a FOSS (Free Open-Source Software) SfM (Structure from Motion) algorithm using different cluster configurations was conducted, leveraging the computing power of ReCaS-Bari data center infrastructure, which hosts several services such as HTC, HPC, IaaS, PaaS. Exploiting the high-computing resources available at clusters and choosing specific set up for the workflow steps, an important reduction of several hours in the processing time was recorded, especially compared to classic photogrammetric programs processed on a single workstation with commercial softwares. The high quality of the image details can be used for land cover classification and preliminary change detection studies using Machine Learning techniques. A subset of the datasets used for the workflow implementation has been considered to test the performance of different Convolutional Neural Networks, using progressively more complex layer sequences, data augmentation and callback functions for training the models. All the results are given in terms of model accuracy and loss and performance evaluation.

Keywords: Photogrammetry, Unmanned Aerial Vehicles, High-Resolution Data, ReCaS-Bari, Deep Neural Networks, Land Cover Classification

Giorgia Miniello, Marco La Salandra

1. Introduction

Photogrammetry [1, 2] is defined as the science of extracting highly reliable three-dimensional spatial information from two-dimensional images. Several commercial workflows are available to perform photogrammetric tasks despite spending a large amount of processing time to complete their steps. In the case of this study, the resulting output of our own algorithm is key to recognize the flooding hazard (through the monitoring of the river conditions, the identification of the channel alterations and morphological changes) and to timely plan the management activities of the emergency after a catastrophic event, with significant time and cost savings. The high performance computing automated photogrammetric workflow fits the scope of direct intervention to safeguard the environment and people's safety, assessing the future scenarios of environmental damage as a function of sudden climate changes.

This document reports the development of a photogrammetric workflow based on Free and Open-Source Software (FOSS) [3], which returns three outputs (the ortophotomosaic, the dense-point cloud and the digital elevation model) managing large amount of data in a reasonable time lapse through the distribution of the most computationally demanding steps on computing clusters hosted by the ReCaS-Bari data center for scientific research. This output can be used for many applications such as territorial mapping and change detection studies of landforms in hydro-geomorphological high-risk areas. Furthermore, the aerial images acquired by Unmanned Aerial Vehicle (UAV) missions can be used for land cover classification. The two main goals of our work are the implementation of an original photogrammetric workflow to process the most computational demanding step of a SfM [4-6] algorithm in the shortest possible amount of time (described in section 2, *Photogrammetric Workflow Implementation*) and the building of a Deep Neural Network for aerial image classification based on the training of an original dataset made of a subset of the aerial images acquired by a drone during two different campaigns (described in section 3, *A Deep Neural Network for Land Cover Classification*). Finally, in section 4 the results and conclusions of our studies are summarized.

2. Photogrammetric Workflow Implementation

As mentioned above, in our study we used a SfM algorithm which aims to fulfill three main steps:

- the detection of key features and tie-points of the images,
- the estimation of calibration parameters and camera positions and orientations,
- the dense point cloud generation.

The data-taking was performed by a "DJI Inspire 2", a quadcopter with aluminium- magnesium composite body and carbon fiber arms equipped with optical sensor "Zenmuse X5S" (20.8 MP, supported lens DJI MFT 15mm/1.7 ASPH, sensor CMOS 4/3", FOV 72°, image resolu- tion 5280 per 3956 pixels), flying at an altitude of about 50 m above the ground level of the take-off location) in order to acquire the desired high-spatial resolution of about 1 cm/pixels. Two reaches of the Basento river near Ferrandina (MT), located in the Basilicata region of southeastern Italy, were explored collecting two datasets. The first one is made of 1139 aerial images (related to 2019), covering an area of approximately 600 per 200 m, and the second is made of 2190 aerial images (related to 2020), covering an area of approximately 1160 per 300 m. Our workflow is based on MicMac, GDAL and Orfeo ToolBox open-source libraries and was developed using two different clusters belonging to the ReCaS-Bari data center, the High-Throughput Computing (HTC) and the High-Performance Computing (HPC) clusters. Firstly, the workflow was implemented on the HTC cluster which counts 128 servers for a total of about 8000 CPU core, with 4GB of RAM per core, and 4PB of parallel disk space. Each computing server, which contains up to 64 slots, can access the whole ReCaS-Bari disk space at a speed of 10 Gbps, using HTCondor as batch system. All the steps of our workflow are summarized in Fig.1. The 1139 image dataset was firstly considered. We found that using a configuration made of 23 worker nodes (WNs) to reach a parallelization of the tasks (*jobs*) for the calculation of tie-points, orthophotos and DEMs on independent subsets of images (50 images per

job), the whole processing time took about 25 hours. A significant improvement was reached using a different cluster configuration, increasing the number WNs (from 23 to 56) and of jobs (reducing the workload on each node from 50 to 3 input images per job). This ensured a more efficient parallel execution limiting at the same time the load on each node of the GPFS file system distributed on the cluster: a significant reduction of the overall processing time was recorded (about 15 hours).
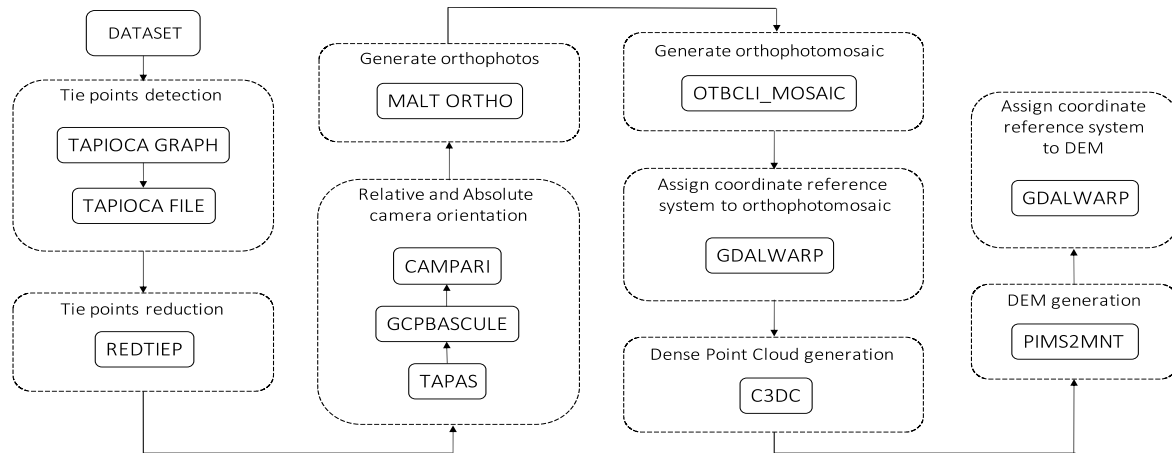
Figure 1. Processing chain of the FOSS photogrammetric workflow

Nevertheless, some considerations could be done in order to manage the main issues related to these configurations. The use of a batch system implies cluster's queue for the job execution, waiting for computing resource availability. Furthermore, several parallel jobs running on a single node heavily affect the workflow time performance since the photogrammetric workflow is characterized by some multi-thread MicMac commands. In order to overcome these issues and improve the results another cluster configuration was set booking a dedicated slot on each WN to run a single job via "pssh", thus ensuring a higher deployment on different nodes and a parallel access to a greater number of nodes. We generated a list of 103 nodes, each one associated with an id number, thus creating a job execution scheme. This brought a not negligible reduction of the whole processing time: ~67% less for the orthophotomosaic generation and ~37% less in the dense point cloud generation compared with the 56-node configuration were recorded. The overall processing time with this configuration is less than 10 hours. Compared to the processing time of the same dataset using a commercial SfM software (Pix4D) on a single workstation, we reached a significant processing time reduction of ~73%. The FOSS photogrammetric workflow was also processed using the second dataset made of 2190 aerial images. In this last case the whole processing time is less than 22 hours. Fig. 2 shows the three outputs of our algorithm for this last dataset: the orthophotmosaic (1.3 cm/pixel), b) the Digital Elevation Model (2.5 cm/pixel) and c) the dense point cloud (~200.000.000 densified points). Finally, both the datasets were used to run the workflow exploiting a single server configuration. The server belonged to the new HPC cluster composed by 5 machines. Each server counts 4 GPU NVIDIA V100 32GB, 96 CPUs, 753.5 GB of RAM and 6 TB SSD Disk. Using The best task configuration, the workflow took less than 4 hours for the 1139 image dataset and less than 10 hours for the 2190 image one to be completed. This result could be considered quite remarkable especially compared with the one obtained using a single server workstation for 1139 aerial images: in this case, the whole processing time was more than 35 hours.

3. A Deep Neural Network for Land Cover Classification

The aerial images acquired by the drone have been used to set up an original dataset that can be divided in classes in order to apply supervised Machine Learning techniques and perform land cover classification studies [7]. Three classes have been considered: "Ground" (817 images), "Vegetation" (1539 images) and "Water" (1198 images). This part of our work is currently in progress to reach the best configuration and results. Different Deep Neural Networks (DNNs) have been tested for territorial classification [8] of the original dataset and only the best two are here presented. We considered 3554 images downgraded to 80 cm/pixel (size of 64 x 48) which is supposed to be close to the spatial resolution of the aerial images captured by a satellite orbiting at very low Earth orbit (such a devise is supposed to be designed for the CLOSE project, see section 4). In the first model, a sequence of pairs of max pooling and convolution layers ending with a dropout layer (30%) and a dense layer was used, setting 150 epochs. A test accuracy of 92.18% was reached although the loss function presented huge fluctuations. In the second model we used a different approach adding a convolutional base of the VGG16 Keras Model (pre-loaded weights) to exploit data augmentation technique (i.e horizontal flip, vertical flip and rotation) and improving the results. In addition to *Model Checkpoint* used in the previous model, the *Early Stopping* and *Reduce LR On Plateau* callback functions were added to limit the overfitting and reduce the learning rate if no improvements are seen after a fixed number of epochs. Using this configuration, the training was early stopped after only 55 epochs, reaching a test accuracy of 94,03% and the test loss was stabilized after less than three decades, as can be observed in Fig. 3.
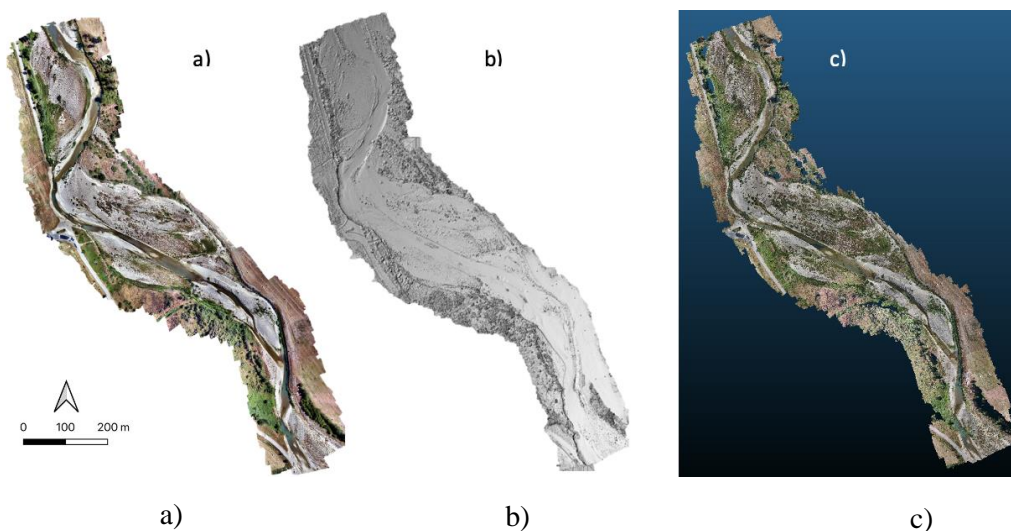

a)   b)   c)

Figure 2. The three outputs of the photogrammetric workflow for the 2190 images dataset. From left to right: a) the orthophotmosaic (1.3 cm/pixel), b) the Digital Elevation Model (2.5 cm/pixel) and c) the dense point cloud (~200.000.000 densified points).

a)   Model loss for VGG16 Model   b)   Model Accuracy for VGG16 Model
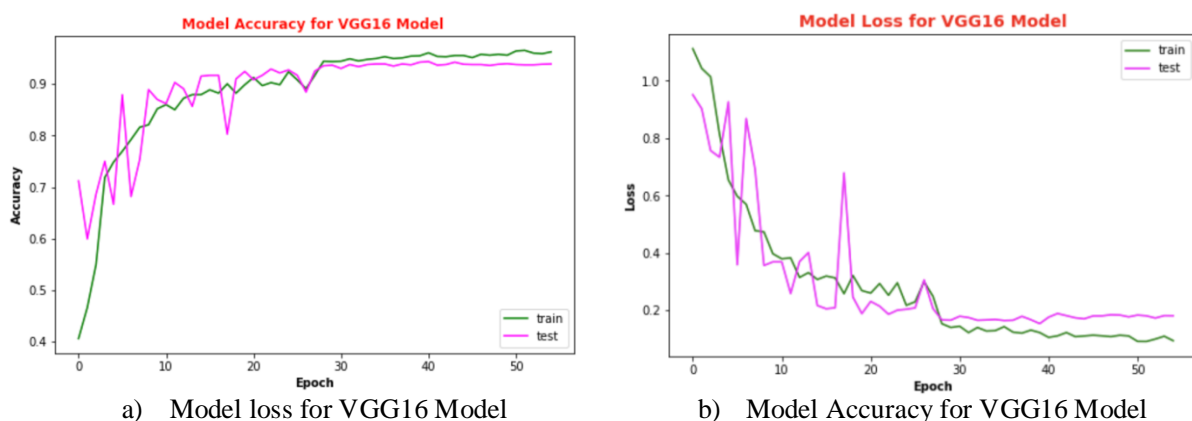
Figure 3. Score for VGG16 DNN model

Even though the number of the classes is limited to three and they are not homogeneously populated, it can be recognized a kind of reciprocity in the misclassification of the classes (e.g., the class "Ground" is mostly misclassified for "Vegetation" and viceversa). As expected, from Fig. 4 it can be observed that the model works better for the most populated dataset and worse for the least one.
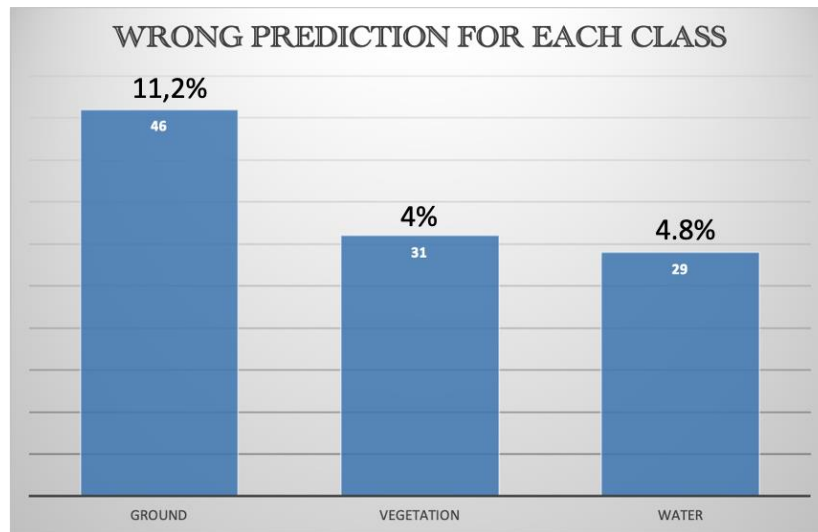

Figure 4. Histogram of wrong predictions for each class

4. Conclusions

An original FOSS photogrammetric workflow to process large dataset of geotagged high-resolution images in a single run was presented. Two original datasets made of 1139 and 2190 high-resolution images (1.09 cm/pixel) respectively have been processed in a relative short time, generating respectively the orthophotomosaic (1.3 cm/pixel), the dense point cloud ($\sim$ 95.000.000 and $\sim$ 200.000.000 densified points) and DEM (2.5 cm/pixel) of the detected areas. Processing time has been optimized distributing the most computationally expensive steps on cluster nodes. A comparison of the processing time using different configurations of computing resources was presented. Results showed that increasing the number of the jobs (thus reducing their workload) and the number of WNs the processing time is drastically reduced. This ensured parallel execution and faster file writing performed by each node on the File System (GPFS). A single-server configuration using the new ReCaS-Bari HPC Cluster was also performed obtaining further improvements with respect to our best result using pssh-configuration. All the outputs obtained are useful to perform detailed hydro-geomorphological analysis of the investigated area. An original dataset of 3554 aerial aerial images has been generated to perform land-cover classification using Machine Learning techniques. We built several DNNs to our original dataset made of images to test 2 different models for land cover classification. Our best model was a VGG16 Keras Model (pre-loaded weights) in which we used data augmentation technique reaching an overall accuracy of $\sim$94% and also reducing test loss to $\sim$15%. This work has been developed in the context of Close to the Earth [9] and RPASInAir [10] projects, Call:"Avviso MIUR n. 1735 del 13/07/2017 AVVISO PER LA PRESENTAZIONE DI PROGETTI DI RICERCA INDUS- TRIALE E SVILUPPO SPERIMENTALE NELLE 12 AREE DI SPECIALIZZAZIONE INDIVIDUATE DAL PNR 2015-2020".

References

[1] Wolf, P.R. and Dewitt, B.A., Elements of Photogrammetry with Applications in GIS, 3rd Edition (2000)

[2] McGlone, J.C. et al., Manual of photogrammetry, American Society for Photogrammetry and Remote Sensing, cop. 2004

[3] Martinez-Rubi, O., Nex, F., Pierrot-Deseilligny, M. et al., 2017, Improving FOSS photogrammetric workflows for processing large image datasets. Open geospatial data, softw. stand. 2, Article number: 12. https://doi.org/10.1186/s40965-017-0024-5

[4] Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, Vol 179: 300–314. https://doi.org/10.1016/j.geomorph.2012.08.021

[5] Ullman S., 1979, The interpretation of structure from motion, Proceedings of the Royal Society of London, Series B Biological Sciences, 203(1153), 405-426.

[6] Wang, X., Rottensteiner, F., Heipke, C., 2019. Structure from motion for ordered and unordered image sets based on random kd forests and global pose estimation, ISPRS Journal of Photogrammetry and Remote Sensing, 147, 19-41: https://doi.org/10.1016/j.isprsjprs.2018.11.009.

[7] Vandana S., 2020, emphLand Cover Classification using Machine Learning Techniques - A Survey, International Journal of Engineering and Technical Research V9(06), DOI - 10.17577/IJERTV9IS060881

[8] K. Simonyan, A. Zisserman, 2015, Very Deep Convolutional Networks for Large-Scale Image Recognition, International Conference on Learning Representations, abs/1409.1556.

[9] https://www.dtascarl.org/progettidta/close-to-the-earth/

[10] https://www.dtascarl.org/progettidta/rpasinair/