# Automatic grain-size curve analyses and unconventional determination of the volume of the muck from TBM through photogrammetry and apple LiDAR sensor

Applied Geomatics (2026) 18:59
https://doi.org/10.1007/s12518-026-00717-y

RESEARCH

# Automatic grain-size curve analyses and unconventional determination of the volume of the muck from TBM through photogrammetry and apple LiDAR sensor

**Andrea Maria Lingua[1] · Francesca Matrone[1] · Francesco Messina[2] · Francesca Parizia[1]**

Received: 19 September 2024 / Accepted: 10 March 2026


**Abstract**

In the context of tunnel excavation with Tunnel Boring Machines (TBM), the possibility of quickly determining, on site and with low-cost tools, both the grain size curve and the volume of the excavated muck, can support the correct setting of the machine and the possible reuse of the material itself for applications such as structural concrete for tunnel linings, road embankments and fillings and so on. For this purpose, unconventional data acquisition methodologies were investigated, such as photogrammetry and LiDAR sensors mounted on smartphones, since they are time-effective from a processing point of view and can be easily adapted and modulated with relatively low budgets. In particular, for the determination of the grain size curve, methods already present in the state-of-the-art were tested, such as BASEGRAIN (expeditiously studied for river pebbles with rounded shapes and not with sharped edges like those of this study), and artificial intelligence techniques for automatic segmentation of rocks from images (Mask R-CNN). For the Volumetric estimation tests were conducted both on the entire sample and on individual grains in order to correctly estimate the volume of the whole excavated sample. The results showed how it is possible to obtain with these techniques on one hand a proper approximation of the granulometric curve, fully comparable with laboratory data, and on the other hand cumulative curves.

**Keywords** TBM · grain size analysis · Granulometric curve · Volume · Muck · Point cloud · iPhone Pro LiDAR · Photogrammetry · Instance segmentation · artificial intelligence · Mask R-CNN

## Introduction

For numerous geological, material and engineering applications, the accurate rock grain quantification is paramount. One essential parameter in this regard is the dimension and the volume of individual rock grains, as it can directly influence aspects such as porosity, permeability, strength, and mineral content of samples or materials to be produced, as well as the possibility of properly setting the Tunnel Boring Machines (TBM) parameters. Traditional methods of volume estimation often involve tedious and destructive techniques or in-lab tests, leading to time-consuming and limited sample sizes with potential alteration of the sample's properties. Photogrammetry or new lightweight and portable sensors, such as smartphones equipped with LiDAR, constitute an innovative and non-invasive approach that leverages the potentialities of the geomatics sector to support the automation of rock grain volume estimation.

Francesca Matrone, Francesco Messina and Francesca Parizia These authors are contributed equally to this work.

✉ Francesco Messina
francesco_messina@polito.it

Andrea Maria Lingua
andrea.lingua@polito.it

Francesca Matrone
francesca.matrone@polito.it

Francesca Parizia
francesca.parizia96@gmail.com

[1] Department of Environment, Land and Infrastructure Engineering (DIATI), Politecnico di Torino, Corso Duca degli Abruzzi, 24, Turin 10129, Italy

[2] Department of Electronics and Telecommunications (DET) - PoliTo Interdepartmental Center for Service Robotics (PIC4SeR), Politecnico di Torino, C.so Francesco Ferrucci, 112, Turin 10138, Italy

Published online: 20 March 2026

Photogrammetry is not a novel approach, but it gained prominence with the latest advancements in digital imaging and computer vision, offering a promising solution to determine the volume of the muck extracted from TBM. By utilizing a series of high-resolution images captured from different angles, photogrammetry can construct three-dimensional models of objects with impressive precision.

The appeal of photogrammetry lies not only in its non-destructive and contactless nature but also in its potential to handle different range of grain shapes and sizes. From angular and irregular mineral grains to smoother sedimentary particles, photogrammetry's adaptability enables researchers to explore volume estimation across various geological contexts. Furthermore, the digital nature of the process allows for seamless data storage, sharing, and analysis, fostering collaboration and time-effective results. However, as with any technique, photogrammetry comes with its challenges and considerations. Factors such as lighting conditions, camera calibration, and image processing algorithms can significantly impact the accuracy of the volume estimation. Moreover, while photogrammetry can handle individual grains effectively, its application to larger rock formations or densely packed samples may require advanced computational approaches, substantial computing resources, and may lead to sub-optimal results. On the other side, the recent advances with the LiDAR sensors, directly mounted on portable devices such as smartphones, can lead to new improvements and a shift in the methodological approach.

This introduction sets the stage for delving into the intricacies of volume estimation of rock grains through geomatics techniques. In this contribution, we thus aim to show an unconventional methodology to estimate the volume of the muck excavated from TBM and the grain size curve. Contrary to what can be found in the state-of-the-art, the main peculiarity of this study is based on the type of rocks analyzed since they possess sharp-edges and very different sizes. This research originates from the needs of a real project where a tunnel is being excavated. In particular, the overall research purposes try to investigate innovative tools to support the logistics of using the excavation materials produced by the Turin-Lyon railway line, for the best sustainability and economic circularity of the process. Given this, the objective of this work is to research the most innovative and reliable methodologies for:

- encourage the reuse of excavated material for applications such as structural concrete for tunnel linings, road embankments and fillings, remodeling (i.e., landfill),
- and to investigate the most automatic and fast approach to determine the petrographic volume directly on-site during tunnel boring, assisting the fast and accurate setting of the TBM.

The research questions we try to answer are, therefore, mainly three:

- which are the best low-cost approaches to be used for the volume and grain size curve determination?
- among these approaches, can photogrammetry be a useful tool to provide a reliable estimation?
- are the newest off-the-shelf sensors as the iPhone Pro LiDAR suitable for these purposes?

The paper structure is the following: Section "Related works" describes the related works, in order to investigate the state-of-the-art 2D and 3D methods for the measurement of particle sizes, highlighting their advantages and disadvantages; Section "Methodology" depicts the data acquisition phases, the sensors used and the data processing; Section "Results" details the results. Finally, we draw the Conclusions in Section "Conclusions" and outline future perspectives.

### Main contributions

The main contributions of this work can be summarized as follows:

- Automated granulometric analysis: Introduction of an AI-based framework leveraging a Mask R-CNN convolutional neural network for automatic segmentation of individual rock fragments from images, enabling the generation of grain size distribution curves comparable to manual and laboratory methods while significantly reducing processing time.
- Use of low-cost and portable sensors: Evaluation of accessible sensing technologies, including digital photogrammetry and the integrated LiDAR sensor of consumer smartphones (Apple iPhone 13 Pro), as practical alternatives to traditional sieve analysis.
- Integrated 2D/3D workflow: Development of a unified pipeline combining orthophoto-based 2D dimensional analysis with 3D point-cloud-based volumetric estimation, including a systematic comparison of volumetric algorithms to identify the most accurate approach.

## Related works

### Determination of particle size: 2D-based methods

The state-of-the-art analysis highlighted the presence of different methodologies to determine the particle size with 2D image data:

- image processing in MATLAB and digital photogrammetry;
- BASEGRAIN software, widely applied in the state-of-the-art;
- PebbleCounts algorithm.

In Alecci (2018), the granulometry of a gravel sample, with clasts of variable dimensions, overlapped and mixed, is analyzed by elaborating images acquired in the laboratory: two methodologies are used, image processing in MATLAB and digital photogrammetry.

In the first methodology using Morphological operators, images are processed to distinguish single grains, and each is approximated to an ellipse (Fig. 1): geometrical properties of each grain (in particular the semi-minor axis) are used to produce the grain size distribution curve.

In the second approach, using digital photogrammetry, the three-dimensional model of the sample is reconstructed, from which level curves are exported and analyzed in MATLAB to separate the belonging grains; this process is carried out using the Watershed algorithm (Vincent and Soille 1991). The grain-size distribution curve derived from image processing showed a trend similar to the laboratory data, unlike the Watershed algorithm, which led to excessive grain segmentation, overestimating the fine and underestimating the coarse.

With reference to our project, elements common to what is studied in Alecci (2018) can be observed in the analysis of heterogeneous and mixed grain sizes, with clasts partially presenting sharp edges and not only rounded edges.

In Detert et al. (2012); Detert and Weitbrecht (2013) was presented a new tool named BASEGRAIN, developed in MATLAB and designed for grains of fluvial origin. It is based on an automated image analysis procedure with a top view, allowing the identification and measurements of grains contained in the represented sample (Fig. 2), estimating the grain size distribution curve in less time than traditional methodologies.

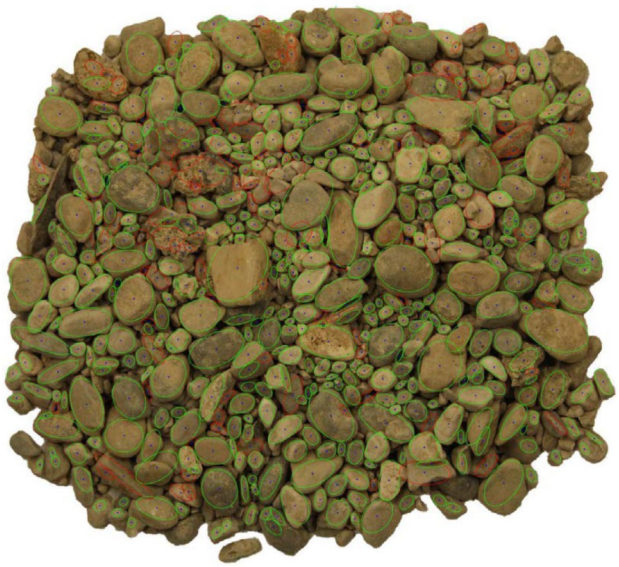


**Fig. 1** Approximation of grains to ellipses, image processing (Alecci 2018)

The authors have implemented the previous methodology proposed by Graham et al. (2005a, 2005b) for river environments, in which, by performing a multi-step processing of digital images, the grain separation is performed with Watershed algorithm, despite the over-segmentation resulting from its application.

BASEGRAIN is widely applied also in applications with UAV (Unmanned Aerial Vehicles), (Astegiano 2015; Langhammer et al. 2017).

In particular, Langhammer et al. performed image-scale calibration to identify the most suitable parameters for subsequent processing stages. Based on the geometric parameters of the detected grains, the grain size distribution curves of interest were then derived; however, similarly to Astegiano (2015), no validation against laboratory data was carried out in their work.

Differently from the previous applications, carried out on debris of fluvial origin, Westoby et al. (2015) have performed a sedimentological characterization of Antarctic moraines, with particles size analysis at a detailed scale, using the BASEGRAIN software. The authors have evaluated the grain size distribution curve of three different portions of soil; to overcome the exclusion of small grains by the software, the grain size distribution curves for diameters beneath 8 mm are produced by Fuller approximation. Finally, the products are compared with curves obtained with traditional method, showing a maximum deviation of 6.3%. In general, the authors observed a slight overestimation of the fine with consequent underestimation of the coarse.

Finally, in Purinton and Bookhagen (2019), PebbleCounts, an open-source algorithm developed in Python, is presented accompanied by a completely automatic method, PebbleCountsAuto. The main purpose, similarly to BASEGRAIN, consists in identifying grains of fluvial origin within digital images.

PebbleCounts is based on the *k-means* algorithm, which is applied in the spatial and spectral domain, in order to separate the single grains: being a semi-automatic procedure, the user is required to select the correctly processed grains, of which the geometric parameters will then be provided, usefully to the evaluation of the granulometry of the entire sample.

Image processing using PebbleCountsAuto consists of automatic identification of grain boundaries and subsequent filtering of incorrectly processed grains. In particular, these

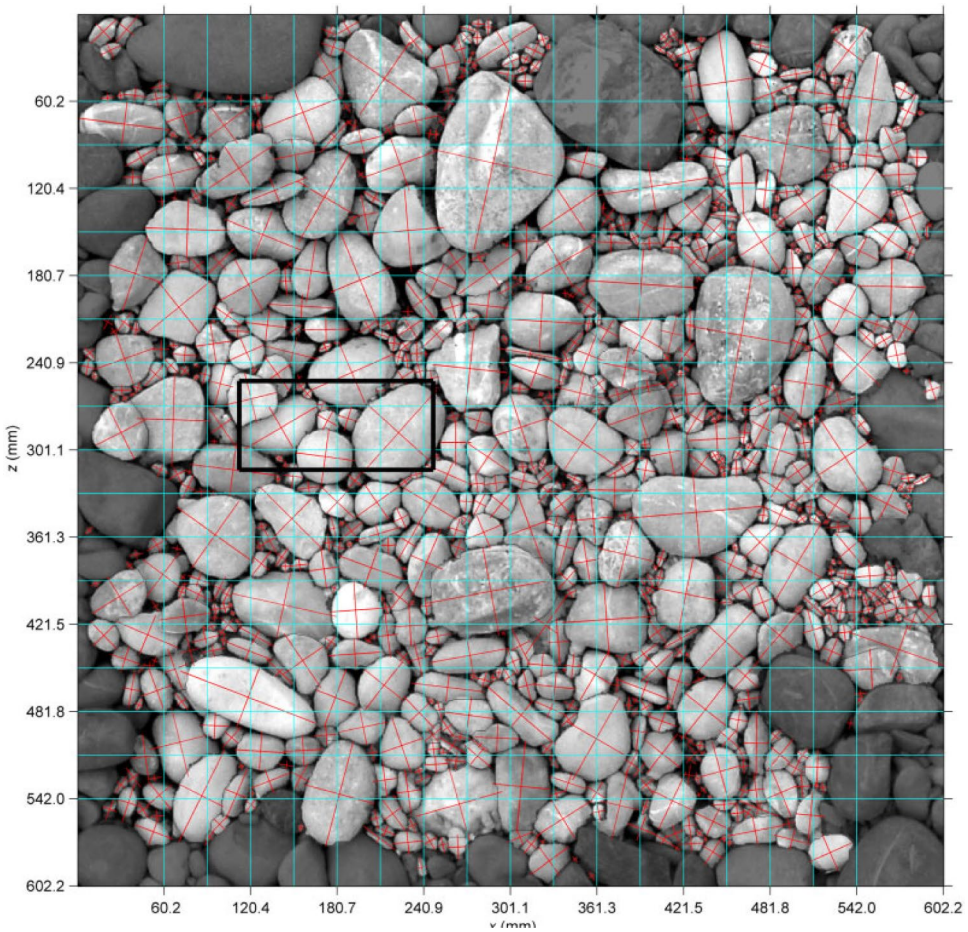



**Fig. 2** Identification of grains with BASEGRAIN software (Detert et al. 2012)

algorithms have been developed by the authors with the aim of processing images depicting very large areas (order of magnitude of hectares), while maintaining high resolution, to the millimeter or centimeter.

The authors have also proposed a comparison with previous methodologies (Watershed algorithm): as shown in Fig. 3. The number of grains identified by PebbleCounts is lower, but with a better precision of delineation of the edges. Similarly, the automatic approach of PebbleCountsAuto is not able to identify the same number of grains traced by the Watershed algorithm, but it increases their accuracy, even if it does not achieve the good results of PebbleCounts.

## Tools and methods for the determination of the drill cored volume - 3D-based methods

The 3D representation and reconstruction of real objects still represent a challenge, despite the ever-new cutting-edge technologies. As Remondino and El-Hakim (2006) explained, it is not an easy task to choose the most appropriate 3D modelling technique in order to meet the generally required elements such as: high geometric accuracy, photorealistic results and full detail modelling, as well as automation, low cost, portability and flexibility of the modelling technique. Nowadays, the generation of a 3D model is mainly achieved using non-contact light-based systems, in particular using active or passive sensors.

The authors summarise 3D modelling techniques in four macro-categories:

- Image-based rendering (IBR)
- Image-based modelling (IBM)
- Range-based modelling
- Combination of image and range-based modelling

The most used techniques today are IBM and range-based. Classically, the IBM methodology uses measurements from 2D images to retrieve 3D information about objects through a mathematical model or obtains 3D data using methods such as Shape from Shading, Shape from Silhouette, Shape from Texture etc. Passive image-based methods for acquiring 3D measurements rely on multiple views, as techniques for extracting three dimensional information from a single image are not yet reliable. In Gilardi et al. (2014), starting from the consideration that rocks of a certain nature having smooth and blunt surfaces can be

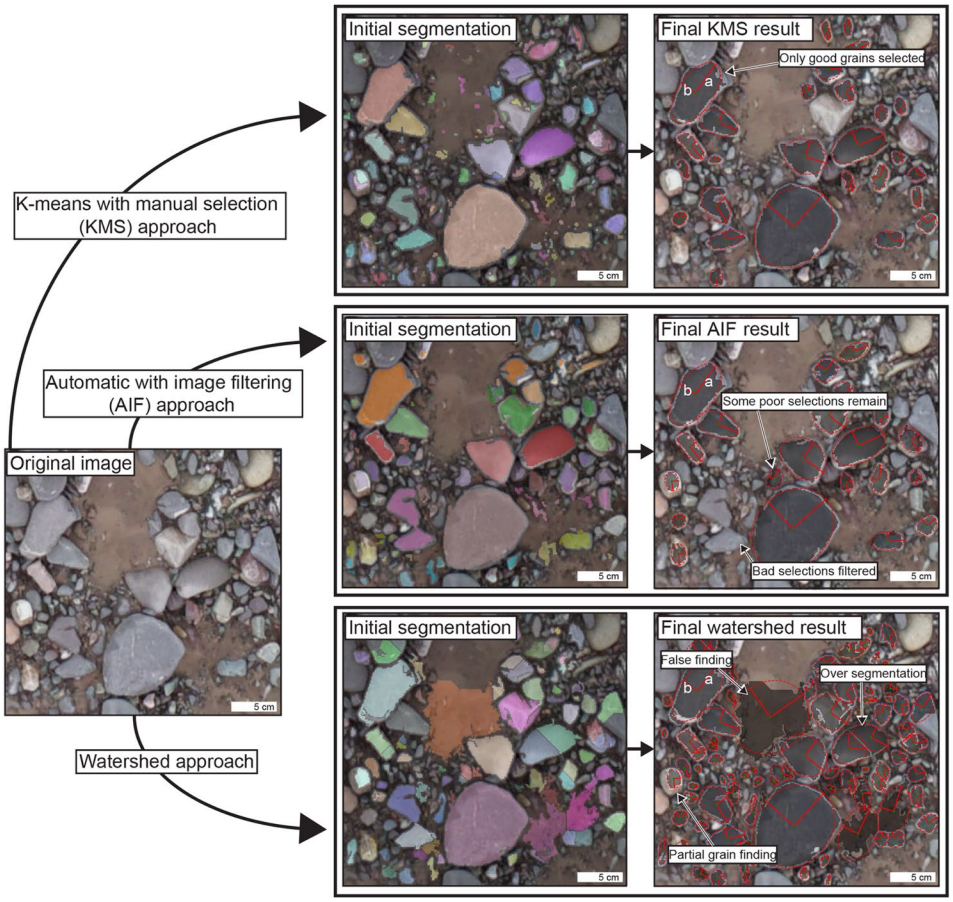



**Fig. 3** Comparison of PebbleCounts, PebbleCountsAuto and the watershed algorithm (Purinton and Bookhagen 2019)

approximated by ellipsoids, they use the following method for 3D reconstruction: from a 2D image, a binary mask is made with the edges of the rocks, then from the equatorial slice of each rock is estimated in polar coordinates, finally using a generalised ellipsoid the mesh of each rock is calculated as shown in Fig. 4.

Although the authors have achieved remarkable results, they report that the average time for reconstructing a single image is 105.421 seconds. In addition, they also report that their method fails to describe the equatorial slices in areas where the edges meet at acute angles, resulting in an oversegmentation of the equatorial rock slice near those areas. In contrast, in Zhao et al. (2020), the authors use multiple images to try to reconstruct the 3D model of the rock. To be precise, they set up an enclosed environment with controlled lighting where the rock to be analysed is placed inside and spun on a turntable, simultaneously acquiring a multitude of images from three cameras positioned at different heights and angles in order to acquire images of the entire rock as shown in Fig. 5.

Agisoft Photoscan (now Agisoft Metashape) processed these sets of images to be aligned and then, using the SfM (Structure from Motion) technique, to obtain a dense point cloud through the block-matching based 3D multi-view dense reconstruction technique. From this dense point cloud, the 3D mesh is reconstructed through Poisson's method with an accuracy level below one millimetre. In this case, one of the major limitations of the method lies in the

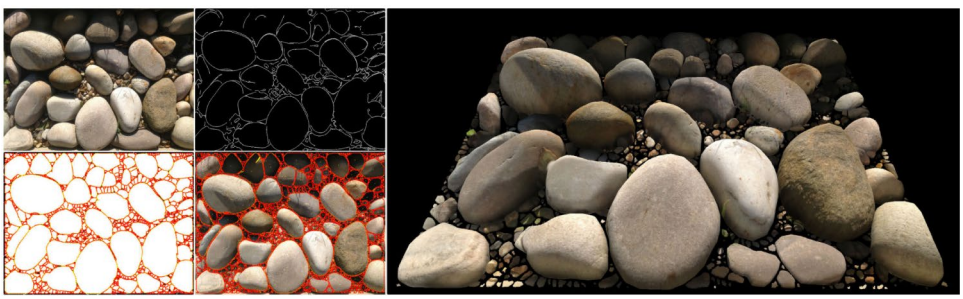


**Fig. 4** Example of 3D reconstruction process from 2D image (Gilardi et al. 2014)

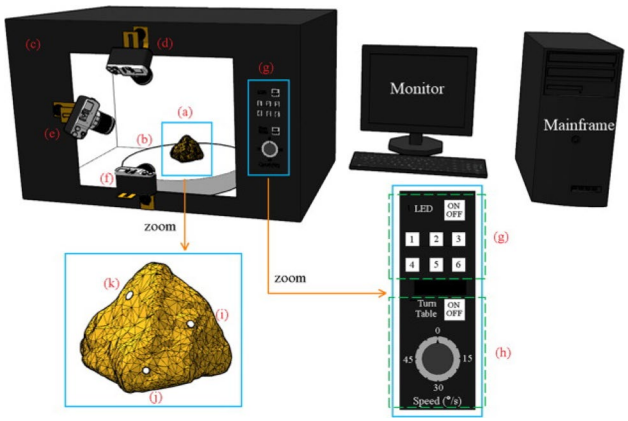



**Fig. 5** Example of 3D reconstruction process from 2D image (Zhao et al. 2020)

fact that only one sample can be analysed at a time, which takes a considerable amount of time if many samples are to be analysed, as in the case of this research. Therefore, this type of approach would be suitable for laboratory analyses instead of *in situ* ones.

Another very interesting example of an IBM technique is presented in Cheung et al. (2005), where the method named Shape-from-Silhouette is used. This technique assumes that the foreground object in an image can be separated from the background. Under this assumption, the original image can be divided into a binary image, foreground and background, which we call a silhouette image. Together with the camera display parameters, the silhouette defines a generalised back-projected cone that contains the real object. The intersection of two cones is called the Visual Hull, which is a bounding geometry of the real 3D object. In this particular work, they propose to use a temporal approach and not the classical spatial one. The difference lies in the fact that in the spatial approach, the number of cameras is increased to obtain many pictures of the object from different angles, while in the temporal approach, only a few cameras are used to take pictures of the object while it is moving. The silhouette information is combined with the pixel colour information from a stereo camera to avoid alignment ambiguities.

For static objects, the problem can be simplified by placing the object on a rotating table so that the movement is known in advance. One of the possible problems is that if the object is not convex, the resulting Visual Hull may not match the real object. On the contrary, as explained in Remondino and El-Hakim (2006), the range-based method directly acquires 3D geometric information of the object. It is based on active sensors and can provide a highly detailed and accurate representation of most shapes. Many of today's commercial solutions are based on triangulation, time-of-flight, continuous wave, interferometry or reflectivity measurement principles. Most of these sensors are very sensitive to the reflective characteristic of the object's surface. Most systems focus only on acquiring 3D geometry, providing only a monochromatic intensity value for each distance value.

The accuracy at certain distances varies significantly from one scanner to another. Furthermore, due to object size, shape and occlusions, it is usually necessary to perform several scans from different positions to cover every part of the object: the alignment and integration of the different scans can affect the final accuracy of the 3D model. In addition, long-range sensors often have problems with edges, resulting in errors or smoothing effects. On the other hand, for small and medium-sized objects, distance-based methods can provide accurate and complete details with a high degree of automation. In Cui et al. (2010), for example, they make use of a ToF camera for the reconstruction of the 3D model of the object. In particular, the authors want to exploit the ability of ToF chambers to measure depth maps at a high frame rate to realise a fast 3D modelling system. The basic problem is that this type of sensors has a classically low resolution, an adverse behaviour to random noise and a significant systematic bias of the measurements that distort the depth map. For this reason, the authors try through an improvement of the specific resolution of the camera and the alignment of the depth maps, by means of a probabilistic alignment algorithm, to obtain a 3D model of satisfying quality. Also in this work, in order to acquire the whole object, a turntable is used or alternatively the camera is rotated around the object.

The fact that in order to obtain a complete model of the object, it needs to rotate on a plate, or the camera needs to rotate around it is quite limiting for several applications.

## Methodology

The goal of this study was a) the automatic definition of the aggregate size (2D), through unconventional visual techniques and an integrated data collection system; b) the investigation and test of new low-cost sensors and faster methods, with respect to the state-of-the.art, to determine a reliable volume of the excavated muck material (3D) (Fig. 6).

### Data acquisitions

The first step in preparing the material coming from the tunnel excavation site was to select and divide it into three samples to perform the standard sieve analysis. One of these three samples (named C2) was finally selected for all the studies described in this paper. The data was acquired with different sensors, a high-performance SONY ILCE-7RM3,

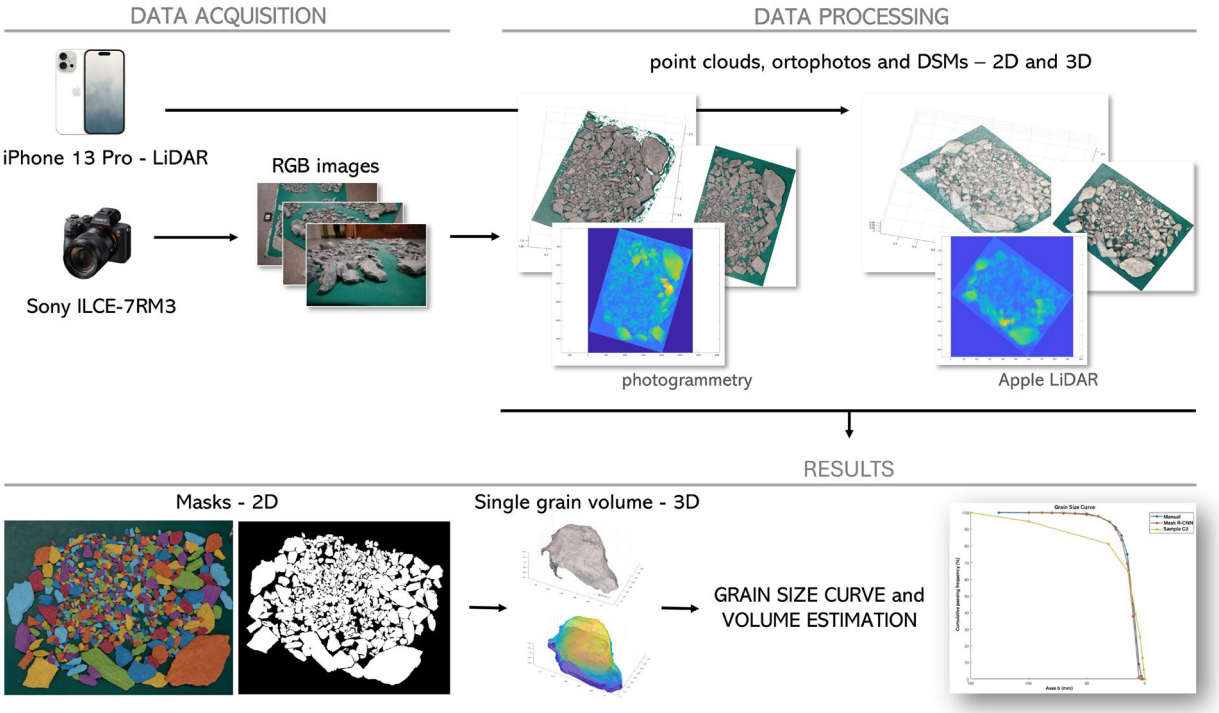



**Fig. 6** Overall workflow from data acquisition and processing to the results

a standard SONY RXII 0 (Table 1) and an iPhone 13 Pro, equipped with a LiDAR sensor to evaluate two low-cost and easy-manageable solutions. As concerns the iPhone LiDAR, only a few official technical specifications are available for the embedded laser sensor. According to Murtiyoso et al. (2021) and Teppati Losè et al. (2022), the sensor is a solid-state LiDAR (SSL) which contains no mechanical motorized parts to ensure scalability and reliability, especially for robotics applications and autonomous vehicle (Wang et al. 2020; García-Gómez et al. 2020). The declared range of the LiDAR sensor is 5 m and the possibility of obtaining a point cloud in nearly real-time has been a determining factor for the choice of this sensor.

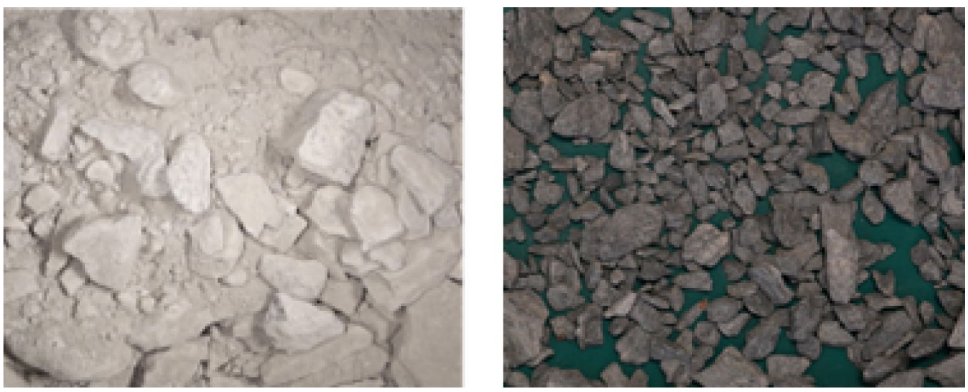
**Table 1** Main spec of SONY ILCE-7RM3, and SONY RXII0

| | SONY ILCE-7RM3 | SONY RX0 II |
|---|---|---|
| Sensor size | 35.9 mm x 24 mm | 1" (13,2 mm x 8,8 mm) |
| Focal length | - | 7,9 mm |
| Photo resolution | 42,4 MP | 15,3 MP |
| Sensor type | CMOS | BSI-CMOS |

## Sample preparation

In the beginning, several rock identification and segmentation attempts were made on the untreated sample. Still, as seen in Fig. 7a, the finer rock particles make this task difficult

**Fig. 7** Excavated material



(a) The material as it is extracted from the TBM

(b) The disposal of the material on a common background.

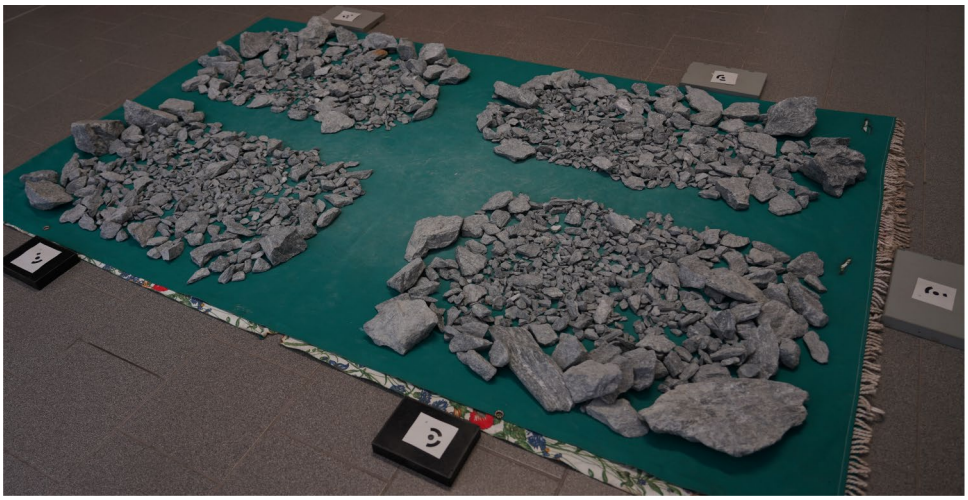


**Fig. 8** The four samples out of the whole testing material

even from a human point of view. Therefore, it was deemed necessary to remove fine particles smaller than 1 cm and subsequently the sample was also washed (for the sake of this type of analyses washing the material is not mandatory).

In the case of in situ operations, these steps could be replaced by a vibrating plate and a rinse water flow or spray.

At the end of this procedure, the rocks were placed on a green cloth to have a contrasting background that simulates the contrast given by the conveyor belt in a real scenario (Fig. 7b).

### Images and LiDAR acquisitions

The selected sample (C2) was further divided into four subsamples, more or less similar, and arranged on the green background as can be seen in Fig. 8.

For the particle size characterization part, a high-resolution nadiral photo was acquired for each subsample with a SONY ILCE-7RM3.

For the volumetric characterization, the first step consisted of placing and surveying five pre-coded targets around the subsamples with a Leica MS 50 total station, using a local reference system and in reflectorless mode. Then, two types of acquisitions were carried out. The first type involved the acquisition of 105 high-resolution (4800x3200) photos around the subsamples, both nadiral and oblique, using a SONY RX0 II. No camera calibration or dedicated lighting setup was performed and adopted, in order to evaluate the validity and potential of the proposed methodology under non-optimal conditions.

The second was a point cloud acquisition performed through the iPhone 13 Pro LiDAR with a circular path around the samples.

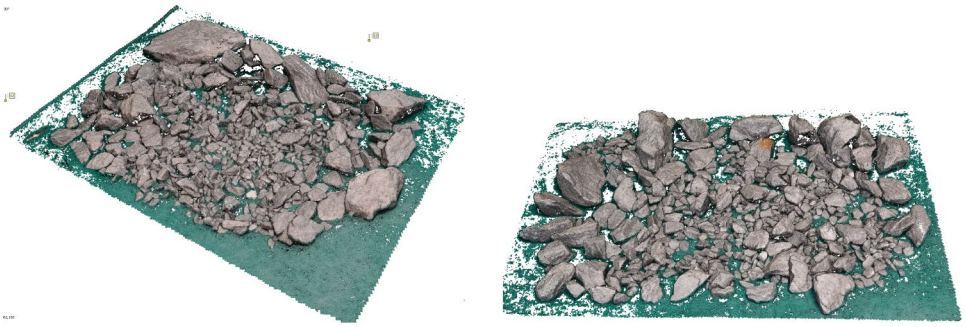
## Data processing

From the first type of acquisitions, a unique point cloud was obtained using the Agisoft Metashape software. The

**Fig. 9** Overall photogrammetric point cloud

measurements acquired with the Leica MS50 total station, operated in reflectorless mode, were used to scale the photogrammetric point cloud. The reported RMSE value of 0.8 mm refers to the internal residuals of the photogrammetric bundle adjustment after scale constraint enforcement, as computed by Agisoft Metashape, and does not represent the absolute metric accuracy of the object scale. Considering the nominal accuracy of reflectorless distance measurements ($2mm + 2ppm$), the achievable absolute scale accuracy is expected to be at the millimetre level. The result is shown in Fig. 9, where the overall point cloud was then split into the four subsamples, each consisting of about 11 million points (Figs. 10a and 10b).

It should be noted that the use of a total station for scaling close-range photogrammetric reconstructions of small objects is not the optimal solution from a metrological standpoint. In similar close-range applications, pre-calibrated bar scales are commonly adopted, as they provide isotropic, stable, and traceable millimetric accuracy. In the present study, the total station was used to reflect a realistic



(a) $Subsample1$. (b) $Subsample2$.

**Fig. 10** Cropped subsamples from the overall photogrammetric point cloud

**Table 2** Comparison of processing results. Rows iPhone 1–4 and Photogrammetric 1–4 report the results for each of the four subsamples processed using the two different methodologies

| Subsamples | Number of points | Processing time | Parameters |
|---|---|---|---|
| iPhone 1 | 293.439 | 1 min | Not possible to set them |
| iPhone 2 | 338.36 | | |
| iPhone 3 | 338.453 | | |
| iPhone 4 | 519.607 | | |
| Photogrammetric 1 | 11.862.182 | 1 hr | Alignment accuracy: High |
| Photogrammetric 2 | 10.446.525 | | Quality: Ultra high |
| Photogrammetric 3 | 11.254.124 | | Filtering mode: Mild |
| Photogrammetric 4 | 11.413.034 | | |

operational scenario in tunnel construction sites, where such instrumentation is often readily available. Nevertheless, future developments of the proposed workflow will include the use of calibrated scale bars to improve absolute accuracy and reduce scale uncertainty.

The processing time required to compute the point cloud was one hour with the following parameters: alignment accuracy set to 'high', the quality set to 'ultra high' and the filtering set to 'mild' (Table 2).

The laptop used for this processing is a DELL XPS15, Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz with 32.0 GB of RAM.

The point clouds acquired with the iPhone (Figs. 11a and 11b) were processed directly by the Scaniverse App, on the iPhone itself.

Generally speaking, the point clouds generated with the iPhone LiDAR have, on one side, a faster acquisition and processing time and, on the other, a lower density than those processed with the photogrammetric technique, as shown in Fig. 12.

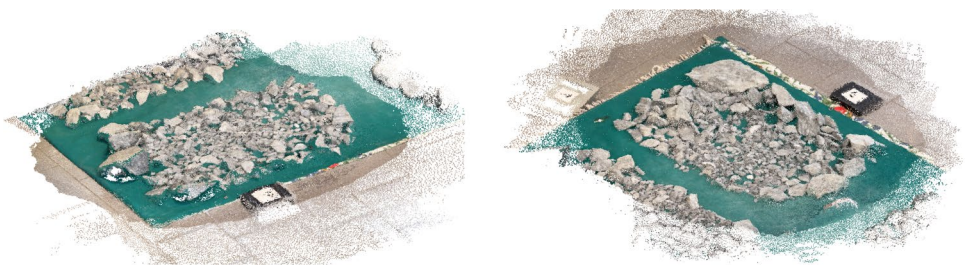
Considering the difference of the density above described, a raster-based elevation model with 0.25 *mm*/*pix* and with an average density of 18 $points/mm^2$ has been finally generated from the DEM (Digital Elevation Model) function in Agisoft Metashape for all four subsamples of the photogrammetric point cloud, as well as the orthophotos, Figure 13.

## Results

### Grain size estimation

In this section we delve into the process of segmentation of single grains through the elaboration of RGB images, trying to obtain a grain-size curve as close as possible to the real one obtained by traditional approaches.

As explained, first attempts were made on untreated samples, without eliminating the finer and dustier components; however, to optimize the particle size recognition, in particular the identification of the edges, it was necessary to remove the fine elements less than 1 *cm*. Employed methods include already existing software and algorithms, such as PebbleCounts, BASEGRAIN and application of a Convolutional Neural Network (Mask R-CNN).



(a) (b)

**Fig. 11** Point clouds acquired with the LiDAR sensor of the iPhone

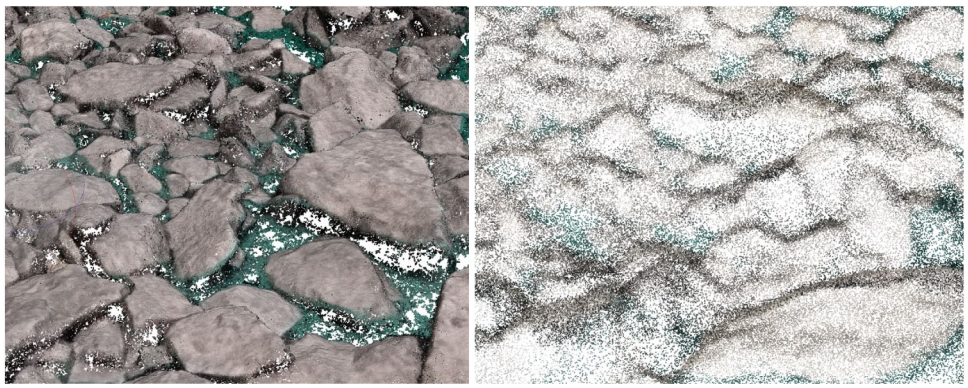


**Fig. 12** Comparison of excavated material in the photogrammetric point cloud (left) and the iPhone (right)

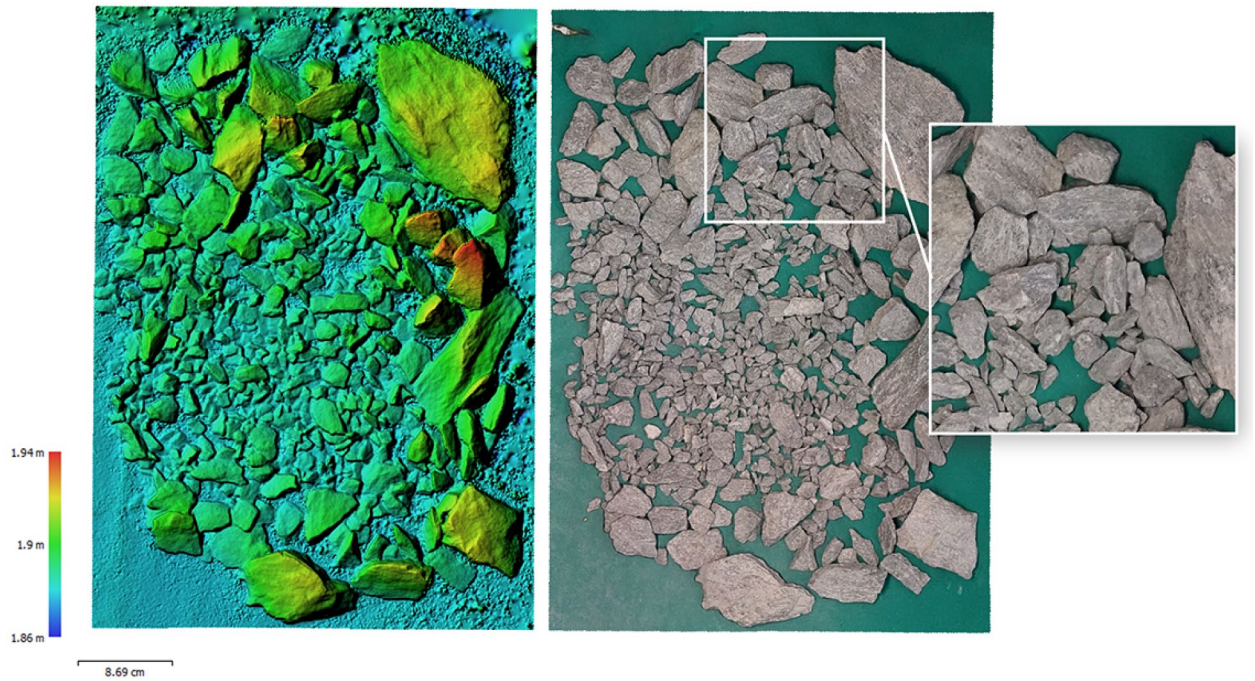



**Fig. 13** Example of raster elevation model (left) and orthophoto (right)

### PebbleCounts and BASEGRAIN

Preliminary tests were conducted on the excavation material as it is (not sieved nor washed) with PebbleCounts and BASEGRAIN to obtain reference results.

In Figure 14, on the left side, an example of the application of PebbleCounts provided by the developers is represented: the identified grains turn out to be very precise; however, a relevant part of the pebbles is not even recognised. Figure 14b on the right shows the results of the abovementioned algorithm on our test material: it is observable that a considerable amount of grains are unrecognized, which led us to consider PebbleCounts not suitable for the case study.

Subsequently, the BASEGRAIN software was tested on samples previously analyzed in the laboratory (Fig. 15). Due to the heterogeneity of the samples, it has not been possible to identify suitable parameters for the correct recognition of the grains, but good results have been obtained after several manual parameter tuning.

Figure 16 shows the comparison between the granulometric curves from laboratory and BASEGRAIN application: the underestimation of the processed curves can be explained by the presence of hidden grains not located on the surface and by the difficult processing of the finer part. To reduce the manual post-processing phase, with consequent reduction of processing times, further attempts have been made on washed and sieved sample 2 (that corresponds to sample C for BASEGRAIN application, and so after this was named sample C2, mentioned in previous chapters), eliminating grains less than 1 cm and the dustier parts to optimize the identification of the grains.

However, any combination of parameters allowed to obtain good results without cases of under-segmentation


**Fig. 14** PebbleCounts on fluvial grains (a) left side and on TELT material (b) right side: failed identification of grains

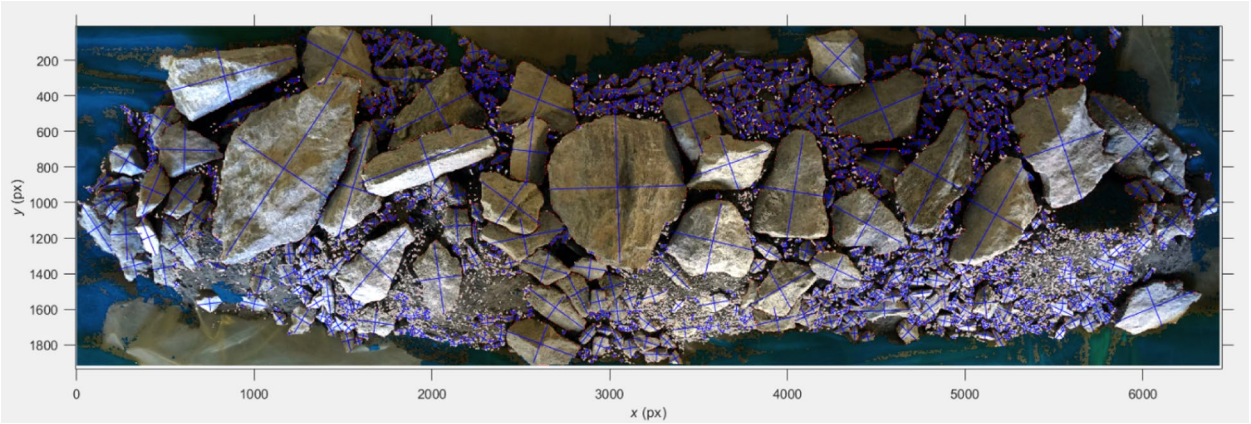



**Fig. 15** Application of BASEGRAIN on not washed and sieved TELT material: good results are obtained only with extensive manual post-processing

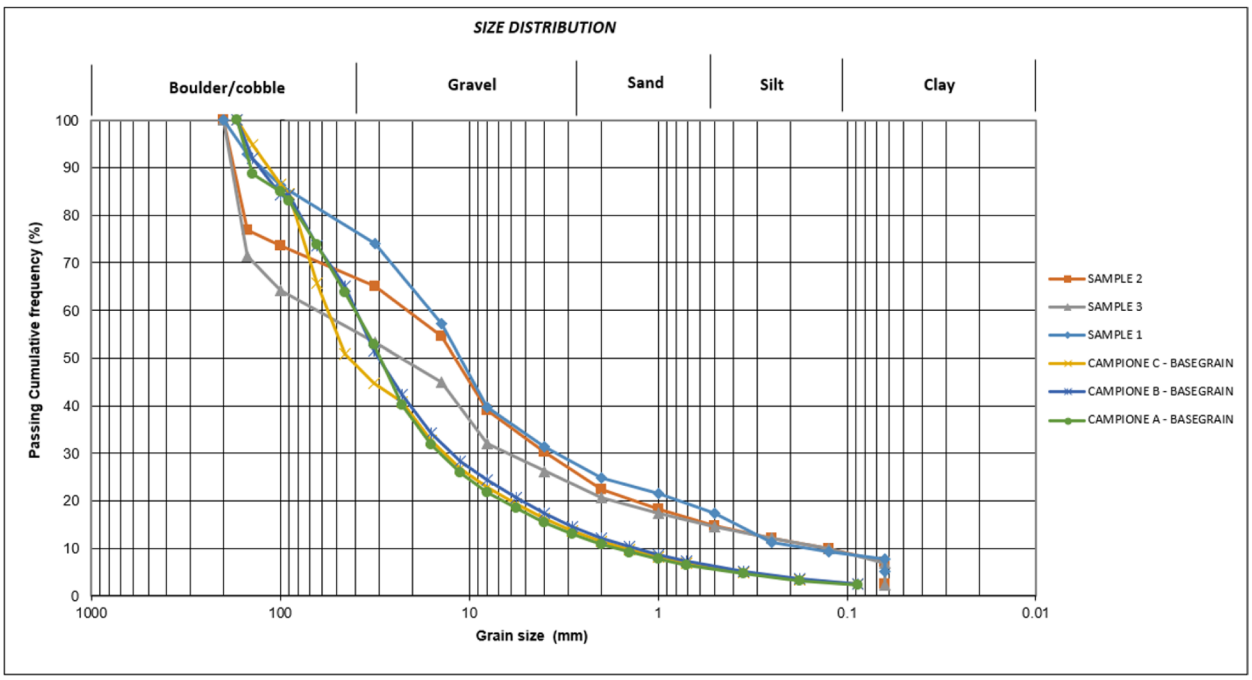



**Fig. 16** Comparison of grain size curves

or over-segmentation (Fig. 17); similarly to Pebble-Counts, BASEGRAIN is also not considered suitable for the case study, in addition to the lack of real-time application.

### Mask R-CNN

The further approach proposed in this paper is using a convolutional neural network to automatically identify the number of elements through the task of *instance segmentation*. Mask R-CNN (He et al. 2017) is a popular deep learning instance segmentation technique that performs pixel-level segmentation on detected objects. The Mask R-CNN algorithm can accommodate multiple classes and overlapping objects. Instance segmentation is an enhanced type of object detection that generates a segmentation map for each detected instance of an object. Instance segmentation treats individual objects as distinct entities, regardless of the class of the objects. In contrast, semantic segmentation treats all objects of the same class as a single entity.

The Mask R-CNN network consists of two stages. The first stage is a region proposal network (RPN), which predicts object proposal bounding boxes based on anchor boxes. The second stage is an R-CNN detector that refines these proposals, classifies them, and computes the pixel-level segmentation for these proposals. The Mask R-CNN model builds on the Faster R-CNN model. Mask R-CNN replaces the ROI max pooling layer in Faster R-CNN with an ROI align layer that provides more accurate sub-pixel level ROI pooling. The Mask R-CNN network also adds a mask branch for pixel-level object segmentation.

About the training phase, a pre-trained network is used on the COCO dataset with the ResNet-50 network as the feature extractor, to which classes of interest (i.e., the Rock class) have been added through the transfer learning technique. The network has been trained with a dataset of about 100 real images, containing only the samples from the excavation, mixed with another 200 images coming from data augmentation techniques. The images needed for the creation of the training, validation and test datasets were obtained from the four high-resolution orthophotographs of each subsample, dividing each of them into patches of 480 x 720 resolution. Then each image was manually labelled in order to obtain the ground-truth masks and bounding boxes.

The performance on the test set can be summarized using two key evaluation metrics: the Mean Average Precision for masks ($mAP_{\text{mask}}$), which reaches 70%, and the mean Intersection over Union (mIoU), which attains 84.7%. These results indicate that the network generates geometrically accurate segmentation masks for correctly detected instances, while the overall instance detection performance is affected by missed detections, primarily due to the strong scale variability between small and large rocks. An example of the segmentation outcome on a complete orthophoto subsample is shown in Fig. 18.

A comparison of the grain size curve results obtained from the sample C2 (sieved with the traditional technique) with the Mask R-CNN segmented samples (obtained from the automatic network segmentation) and the manually labelled ground truth is shown in Fig. 19. As can be appreciated from the graph, the curve obtained from this methodology well approximates the one manually obtained, except for rocks with b-axis greater than 100 mm. This error is probably due to training dataset bias, which probably originates from the division of the original image into smaller patches that results in a possible subdivision of larger rocks. Moreover, in some patches, there was only one rock occupying more than 70% of the image. Since these cases were rare compared to the number of input photos, it was probably not enough for the network to learn and understand how to correctly segment under these conditions.

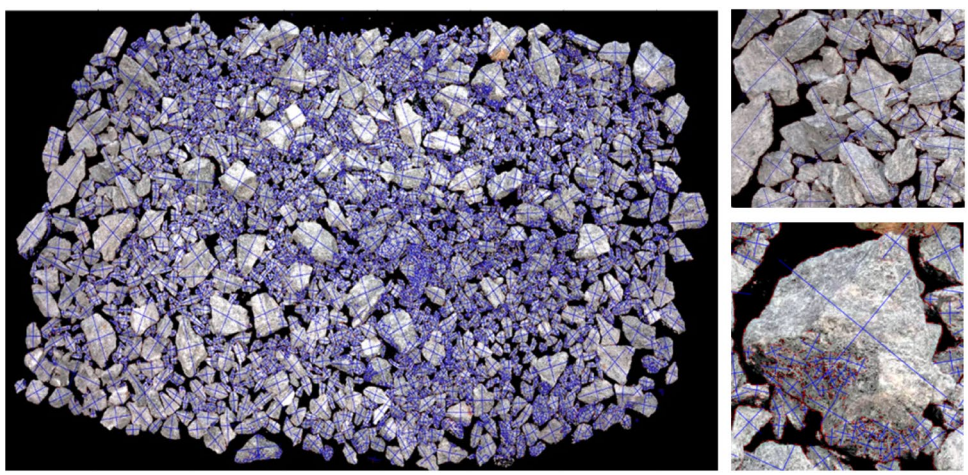


**Fig. 17** Application of BASEGRAIN on treated TELT material (sample C2): problems of under-segmentation (top right) and over-segmentation (bottom right)

**Fig. 18** Manually labelled ground truth (left), obtained results using Mask R-CNN on a subsample orthophoto (center-right) and rock detection detail on a single patch (right)

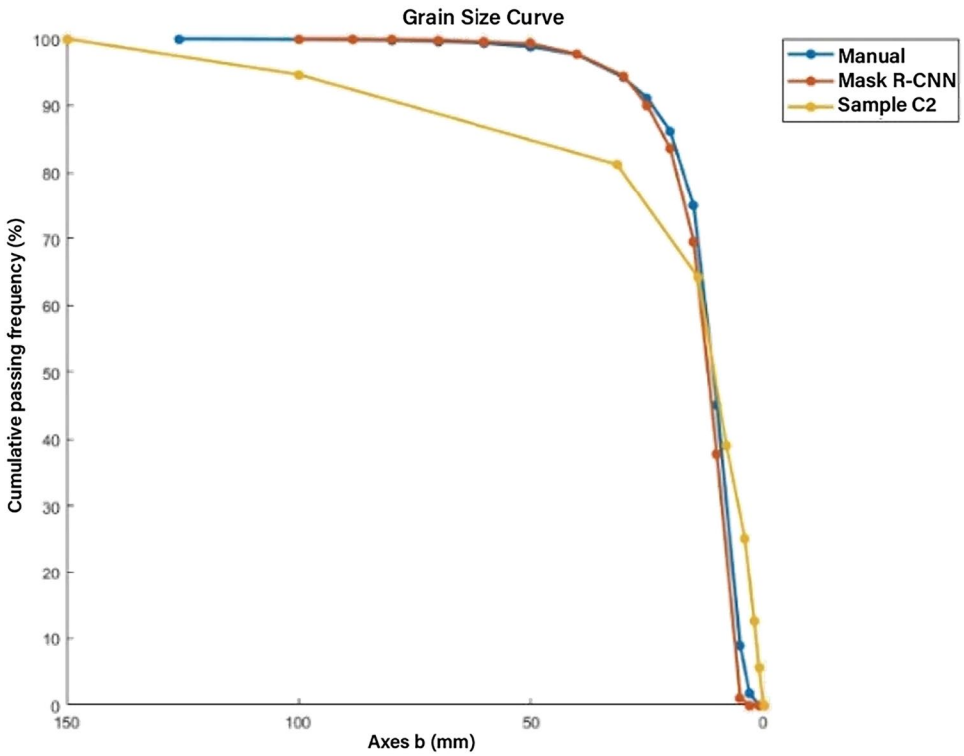



**Fig. 19** Comparison between grain size curves obtained through manual labelling or the Mask R-CNN

## Volumetric estimation

Once the grain size analysis was completed, our focus shifted to volumetric estimation of the excavated material, starting from the previously acquired point clouds. Also in this case, the analysis concerned sample C2 that has been divided into four parts, whose results have then been summed together and compared with known laboratory data (sample C2 has a mass of 46170 g and a density of 2.7 $g/cm^3$: the volume is equal to 0.0171 $m^3$)

The total volume of the sample has been estimated by applying different techniques, implemented for both photogrammetric and LiDAR data:

- evaluating the complessive volume of each subsample in MATLAB environment and CloudCompare;
- extrapolating single grains and evaluating their volume in MATLAB.

### Entire sample analysis

In *CloudCompare*, the *'2.5D Volume'* tool has been used, which allows for evaluating the volume of the point cloud on the basis of the distance of each point from a reference height, which is set equal to the minimum z coordinate of points belonging to the green floor. By setting the size of a reference grid (step), a raster is produced, in which the distance value from the imposed base to each cell is associated. An example of the result obtained for one of the subsamples in *CloudCompare* is shown in Fig. 20. In Table 3 the results obtained for each subsample are summarized.

In *MATLAB*, a different kind of process was followed for the volume estimation. Starting from the orthophoto, a mask of the entire group of rocks is produced using morphological operations, Figure 21. Then, only the pixels of the raster-based elevation model belonging to this mask have been isolated, therefore excluding those not representative

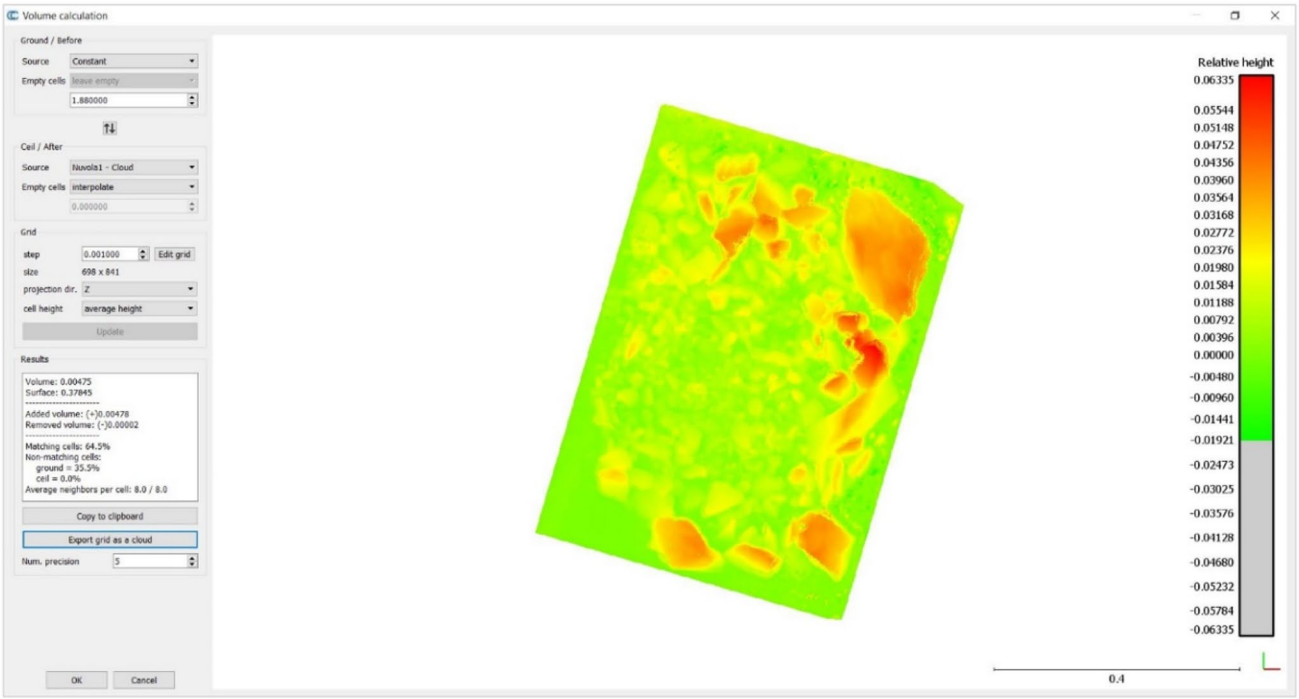



**Fig. 20** Volume estimate in CloudCompare: part 1 of Sample C2

**Table 3** Results of volume estimation for the entire sample C2

| | | PHOTOGRAMMETRY ($m^3$) | | LiDAR ($m^3$) | |
|---|---|---|---|---|---|
| | | CloudCompare | MATLAB | CloudCompare | MATLAB |
| SAMPLE C 2 | PART 1 | 0.0048 | 0.0075 | 0.0034 | 0.0053 |
| | PART 2 | 0.0046 | 0.0054 | 0.0040 | 0.0055 |
| | PART 3 | 0.0041 | 0.0041 | 0.0035 | 0.0050 |
| | PART 4 | 0.0034 | 0.0049 | 0.0030 | 0.0047 |
| | **V ($m^3$)** | **0.0169** | **0.0219** | **0.0138** | **0.0205** |
| | **Δ ($m^3$)** | **-0.0002** | **0.0048** | **-0.0033** | **0.0034** |
| | **Δ (%)** | **-1.35** | **27.92** | **-19.06** | **19.89** |

of the rocks. Similarly to CloudCompare, for each of these cells, a volume has been associated, knowing its area and associating a height given by the difference between the raster value in that cell and a reference value (ground), then all the partial volumes are summed.

As can been seen in Table 3, the best results are obtained by applying the CloudCompare '2.5D Volume' method on the photogrammetric point clouds, with a percentage deviation from the laboratory value of -1,35%. In the case of the LiDAR point cloud, there appears to be a not-too-marked underestimation (-19,06%), probably due to the lower density of the point clouds. On the contrary, the MATLAB methodology, exploiting the raster model, appears to cause overestimation of the volume for both the point clouds (+27,92% and +19,89%). The symbol Δ denotes the difference between the results obtained with each column-based method, as reported in Table 3, and the ground truth values of sample C2 derived from standard sieving analysis. The same definition applies to Tables 5 and 6.

### Single grains analysis

The methodology for the volume estimation of the single grains is developed in MATLAB. In this case, the point clouds and masks coming from the Instance Segmentation network have been used. First of all, each mask obtained from the automatic segmentation has been re-projected onto the orthophotos (Fig. 22 b), to obtain a correspondence between the points of the rocks in the masks and the relative points belonging to both types of point clouds

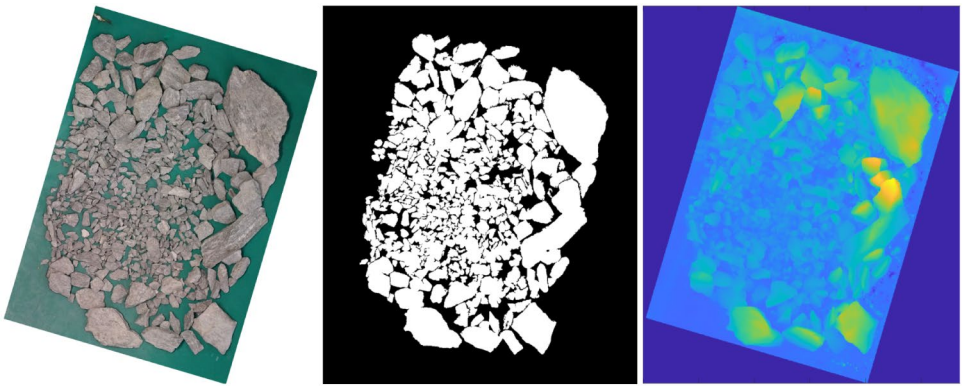


**Fig. 21** Orthophoto, mask and raster elevation model of part 1 of sample C2, MATLAB

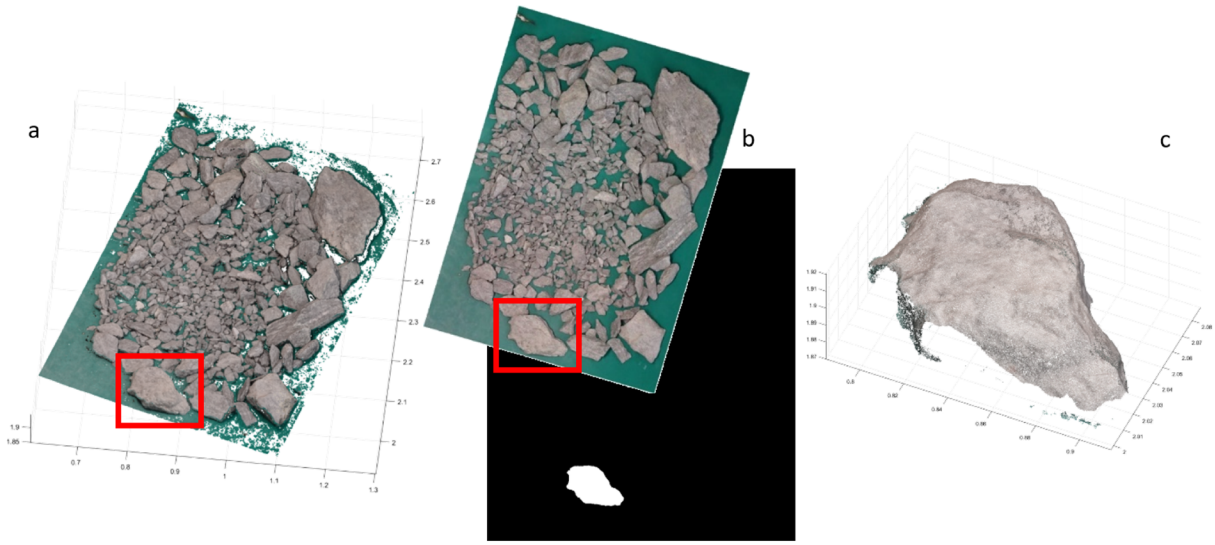



**Fig. 22** Extrapolation of single rock point cloud (c) starting from original point cloud (a) and mask (b)

(photogrammetric and Lidar ones). An example of single grain extraction from the point cloud is shown in Fig. 22 c.

Point clouds of individual rocks have been further processed: a base has been added to each point cloud, projecting each point at a fixed $z$, equal to the minimum $z$ of the rock itself; an example of this process can be seen in Fig. 23. This step has been carried out in order to avoid possible underestimations of the volume, maybe caused by occluded parts of rocks.

Once isolated the individual rocks, their volume is calculated by the DSM method and by applying two MATLAB functions, *'Convhull'* and '*Boundary'*, which both operate directly on the point cloud, enclosing it in a closed surface and evaluating its volume; since the two functions generate different volumes for the same points, they have been compared.

Further geometric information of rocks have also been evaluated:

- the difference between the maximum z and the minimum z has been used to evaluate the third dimension of the rock;
- starting from the masks and approximating the rock to an ellipse, the a and b axes have been measured.

Table 4 resumes some of these results for several rocks. Volume estimation from single grains, photogrammetry.

The volume of the whole sample has then been calculated, for each technique applied, as the sum of all the individual volumes of the rocks. The three methods used (*DSM, Convhull* and *Boundary*), have been applied both for the photogrammetric (Table 5) and LiDAR point clouds (Table 6), and they were put in comparison with the manually segmented masks.

Based on the results obtained from the photogrammetric data, in the case of manually segmented masks, we can observe that the Convhull method turns out to be the one with the lowest percentage deviation compared to the results obtained from the analysis done in the laboratory. While as for the results obtained using the automatically segmented masks we can see that the percentage deviation again using the Convhull function shows a significant underestimation

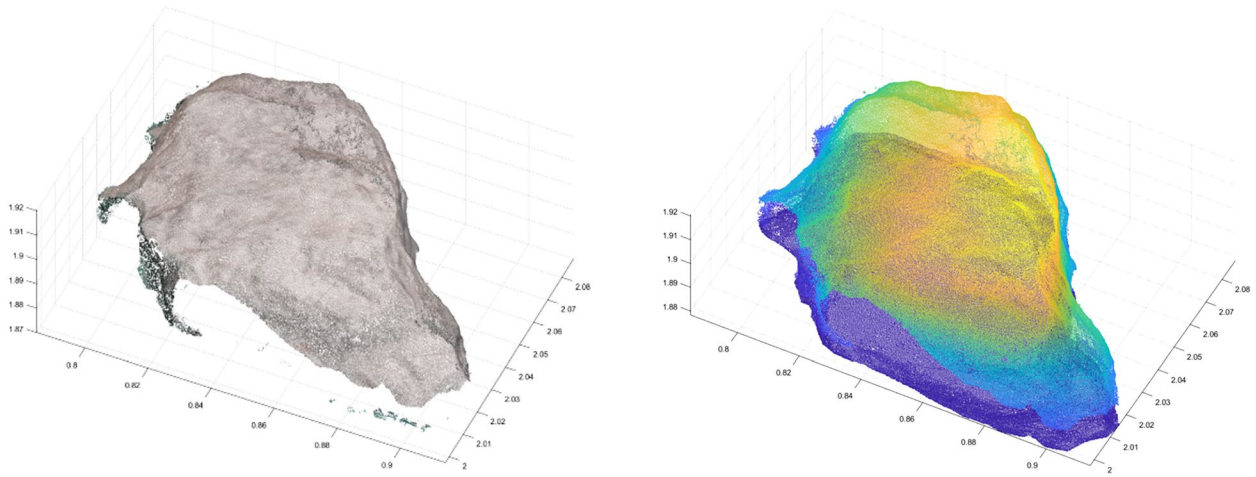


**Fig. 23** Adding base points to the rock point cloud

**Table 4** Geometric parameters and volume of single grains

| | Axis a (cm) | Axis b (cm) | Axis c (cm) | $V_i$ - DSM ($cm^3$) | $V_i$ - Convhull ($cm^3$) | $V_i$ - Boundary ($cm^3$) |
|---|---|---|---|---|---|---|
| **Rock 1** | 2.10 | 1.30 | 0.92 | 4.56 | 1.38 | 1.09 |
| **Rock 2** | 7.15 | 4.31 | 3.09 | 77.19 | 53.85 | 51.90 |
| **Rock 3** | 5.36 | 3.59 | 2.50 | 47.95 | 26.93 | 24.32 |
| **Rock 4** | 2.02 | 1.03 | 0.29 | 3.16 | 0.36 | 0.32 |
| **Rock 5** | 1.01 | 0.64 | 0.45 | 0.98 | 0.12 | 0.11 |
| **Rock 6** | 2.45 | 1.12 | 1.18 | 4.67 | 1.65 | 1.57 |
| **Rock 7** | 1.43 | 0.77 | 0.49 | 1.65 | 0.35 | 0.29 |
| **Rock 8** | 1.79 | 0.98 | 0.44 | 2.64 | 0.37 | 0.23 |
| **Rock 9** | 21.47 | 12.56 | 4.93 | 1056.52 | 845.14 | 784.97 |
| **Rock 10** | 10.52 | 8.23 | 4.47 | 331.51 | 261.00 | 244.02 |
| **Rock 2562** | 1.71 | 0.32 | 0.74 | 0.61 | 0.22 | 0.11 |
| **Rock 2563** | 1.42 | 0.91 | 0.65 | 1.82 | 0.58 | 0.32 |
| **Rock 2564** | 2.01 | 1.02 | 0.42 | 2.08 | 0.63 | 0.25 |
| **Rock 2565** | 1.00 | 0.36 | 0.37 | 0.40 | 0.06 | 0.05 |
| **Rock 2566** | 1.05 | 0.66 | 0.41 | 0.80 | 0.15 | 0.12 |

**Table 5** Volume estimation from single grains, photogrammetry

| | MANUAL MASKS | | | INSTANCE SEGMENTATION | | |
|---|---|---|---|---|---|---|
| | $\Sigma V_i$, DSM | $\Sigma V_i$, Convhull | $\Sigma V_i$, Boundary | $\Sigma V_i$, DSM | $\Sigma V_i$, Convhull | $\Sigma V_i$, Boundary |
| **V ($m^3$)** | **0.0237** | **0.0158** | **0.0140** | **0.0180** | **0.0108** | **0.0097** |
| **Δ ($m^3$)** | **0.0066** | **-0.0013** | **-0.0031** | **0.0009** | **-0.0063** | **-0.0074** |
| **Δ (%)** | **38.39** | **-7.38** | **-18.11** | **5.10** | **-36.72** | **-43.06** |

**Table 6** Volume estimation from single grains, LiDAR

| | MANUAL MASKS | | | INSTANCE SEGMENTATION | | |
|---|---|---|---|---|---|---|
| | $\Sigma V_i$, DSM | $\Sigma V_i$, Convhull | $\Sigma V_i$, Boundary | $\Sigma V_i$, DSM | $\Sigma V_i$, Convhull | $\Sigma V_i$, Boundary |
| **V ($m^3$)** | **0.0207** | **0.0000** | **0.0072** | **0.0160** | **0.0056** | **0.0050** |
| **Δ ($m^3$)** | **0.0036** | **-0.0171** | **-0.0099** | **-0.0011** | **-0.0115** | **-0.0121** |
| **Δ (%)** | **21.05** | **-100.00** | **-57.91** | **-6.27** | **-67.32** | **-70.79** |

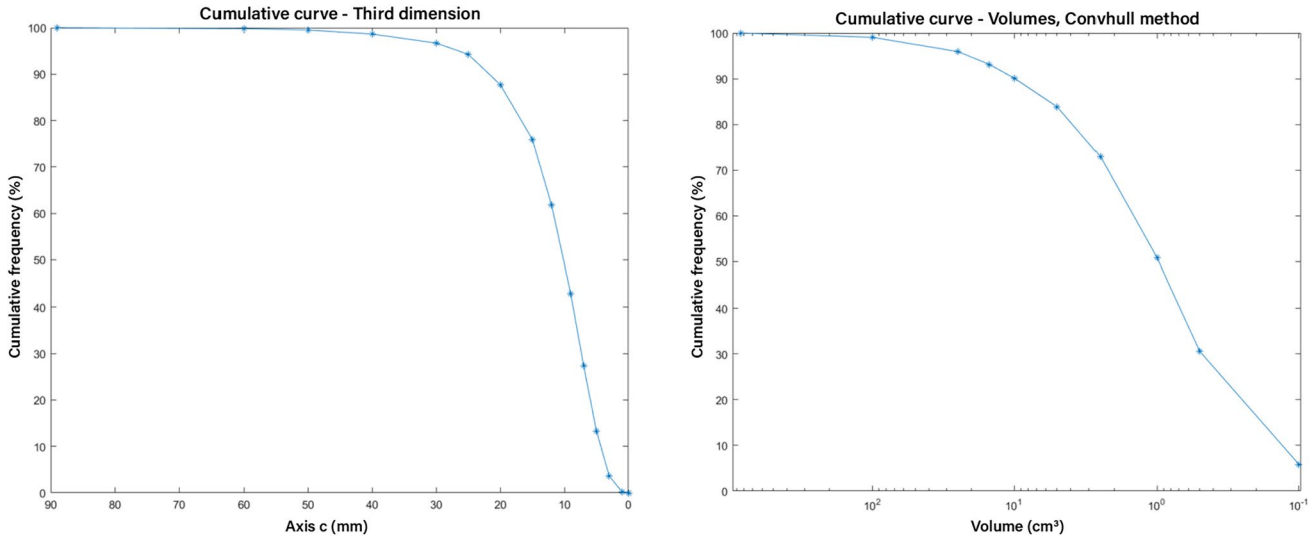



**Fig. 24** Cumulative curves: third dimension (left) and volume estimate by Convhull (right)

of the volume that is attributable in good part to the under-segmentation resulting from the network (that we have already underlined in Section "Mask R-CNN").

Instead, the approach based on DSM seems to generate, like in the previous case (Table 3 – MATLAB column), an overestimation of the volume, which in the case of Instance Segmentation masks is probably balanced with the masks under-segmentation, resulting in a final volume closer to that of laboratory.

The limitations observed in the volumetric analysis based on iPhone LiDAR data are primarily related to sensor accuracy, point cloud density, and scale effects. Commercial solid-state LiDAR sensors embedded in smartphones are affected by both random noise and systematic depth errors, which may introduce local surface smoothing and bias in the reconstructed geometry. The acquired LiDAR point clouds exhibit an average point spacing on the order of a few millimetres (approximately 1–3 mm), whereas the photogrammetric point cloud is more than an order of magnitude denser. By taking a look at the Table 6, it can be seen that the low density, the depth noise and roughness of the point cloud do not allow any method to obtain acceptable results: for both Boundary and Convhull algorithms, there are strong underestimations. As before, the DSM possible overestimations looks balanced from the 2D under-segmentation.

Cumulative curves referring to axis a, axis b, third dimension and calculated volumes (DSM, Convhull and Boundary) are also produced: two examples are depicted in Fig. 24. These curves are not comparable with laboratory data, but could be useful for the optimization of the excavation through TBM, knowing the single grains distribution and not only the entire sample volume.

## Conclusions

Two possible methodologies for particle size analysis of the excavated material were presented: software applications such as PebbleCounts and BASEGRAIN, and AI-based convolutional neural networks (Mask R-CNN). Recalling the initial research questions, BASEGRAIN and PebbleCounts, developed for grains of fluvial origin, turned out to be not suitable for the purpose, requiring extensive manual correction and post-process adjustments. The neural network, on the other hand, was found to be more appropriate, although the segmentation results obtained with Mask R-CNN are better than those produced by software like BASEGRAIN and PebbleCounts, the presented results are still not applicable to the actual construction site since almost perfect recognition of the grains has not yet been achieved. In fact, in the case of large grains, there is a definite underestimation, also caused by the way the patches are created, which splits the majority of bigger rocks, introducing a bias in the training dataset. Despite this limitation, the comparison between the produced grain size curve and that evaluated in laboratory by traditional techniques can be considered satisfactory for a first approximate estimation that can be automated directly on site. Moreover, the choice of the CNN allows the optimisation of the time for analysis and processing, the network takes less than 1 second to segment an image of 480 x 720, using a laptop with the following characteristics: LAPTOP-4GVIAO08, Intel(R) Core(TM) i7-7700HQ CPU @ 2.80GHz 2.81 GHz with 16.0 GB of RAM. As for the convolutional neural network, there is still much room for improvement in segmentation results. Some methods that could be applied in future research include:

- an expansion of the training dataset, maybe using different samples coming from different excavation sites or using different TBM excavation parameter.
- the depth data coming from the DSM could be used jointly with the RGB images to make more robust predictions.
- after the reconstruction of the orthomosaic with the segmented patches, in order to reduce border effects and try to recompose some of the bigger rocks, as described above, some morphological operations could be used.

About the volumetric evaluation, the analysis was carried out either on the entire sample or by analyzing the individual grains separately, always using both LiDAR and photogrammetric data: methods that have provided results much closer to the known laboratory value were found to be the '2.5D Volume', applied to the entire mass in CloudCompare, and the 'Convhull' algorithm, used on individual rocks in MATLAB, both on photogrammetric data. The latter, however, derives from the application of manually drawn masks, and not from the previous instance segmentation: so, it can be considered only a potential method of volumetric evaluation, in case of optimization of the previous phase, in order not to underestimate the total volume. The analysis of individual grains would also allow the optimization of TBM excavation parameters. As regards the third research question, the achieved results showed that the tested low-cost iPhone LiDAR sensor showed limitations in volumetric estimation for the analyzed TBM muck samples, mainly due to point cloud roughness, noise, scale-dependent resolution, and systematic depth errors. In particular, these point cloud features were not sufficient to reliably resolve smaller grains and sharp edges, resulting in surface smoothing and volume underestimation. These results do not imply that smartphone LiDAR technology is generally unsuitable for volumetric analysis, but rather that its effectiveness strongly depends on the relationship between sensor resolution and object scale. Future investigations using higher-performing LiDAR sensors or hybrid approaches integrating photogrammetric data may overcome these limitations. Finally, the system developed and tested here should be adequately redesigned for construction site activities with a conveyor belt and sensors protected from excessive dust and with perfectly absorbent side panels, to avoid unwanted light reflections.

In summary, this work demonstrates the feasibility of an automated, AI-driven granulometric analysis framework that combines image-based rock segmentation with low-cost, portable sensing technologies. By integrating digital photogrammetry and consumer-grade smartphone LiDAR within a unified 2D/3D workflow, the proposed approach enables accurate estimation of particle size distributions and volumes while significantly reducing processing time and operational complexity compared to traditional sieve-based methods. The results confirm that accessible sensing solutions, when coupled with advanced computer vision techniques, can provide reliable and scalable alternatives for granulometric characterization in real-world industrial and field applications.

**Acknowledgements** The authors would like to acknowledge the TELT SAS Society, which funded the project, and the interdepartmental centre SISCON "Safety of InfraStructures and CONstructions" of Politecnico di Torino, which was in charge of the research under the scientific direction and supervision of Prof. Alessandro Fantilli.

**Author Contributions** All authors contributed equally to this work. A.M.L. supervised the project and coordinated the research activity. A.M.L. and F.Messina developed the concept and designed experiments. F.Matrone, F.Messina and F.P. performed the experiments, collected and analysed the data. F.Messina and F.P. wrote the code. F.Matrone and F.Messina wrote the manuscript.

**Funding** Open access funding provided by Politecnico di Torino within the CRUI-CARE Agreement.

**Data Availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing interests** The authors declare no competing interests.



## References

Alecci C (2018) Analisi fotogrammetrica dei terreni: applicazione, limiti e differenze con la granulometria classica

Astegiano L (2015) Utilizzo di piattaforme apr per il monitoraggio ambientale dei corsi d'acqua

Cheung K-M, Baker S, Kanade T (2005) Shape-from-silhouette across time part i: Theory and algorithms. Int J Comput Vision 62:221–247

Cui Y, Schuon S, Chan D, Thrun S, Theobalt C (2010) 3d shape scanning with a time-of-flight camera. In: 2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, pp 1173–1180

Detert M, Weitbrecht V (2013) User guide to gravelometric image analysis by basegrain. In: Fukuoka S, Nakagawa H, Sumi T, Zhang H (eds) Advances in science and research, pp 1789–1795

Detert M, Weitbrecht V et al (2012) Automatic object detection to analyze the geometry of gravel grains-a free stand-alone tool. In: River Flow. Taylor & Francis Group London, UK, vol 2012, pp 595–600

García-Gómez P, Royo S, Rodrigo N, Casas JR (2020) Geometric model and calibration method for a solid-state lidar. Sensors 20(10):2898

Gilardi M, Watten PL, Newbury PF (2014) Unsupervised three-dimensional reconstruction of small rocks from a single two-dimensional image. In: Eurographics (Short Papers), pp 29–32

Graham DJ, Rice SP, Reid I (2005) A transferable method for the automated grain sizing of river gravels. Water Resourc Res 41(7)

Graham DJ, Reid I, Rice SP (2005) Automated sizing of coarse-grained sediments: image-processing procedures. Math Geol 37:1–28

He K, Gkioxari G, Dollár P, Girshick R (2017) Mask R-CNN. In: Proceedings of the IEEE international conference on computer vision, pp 2961–2969

Langhammer J, Lendzioch T, Miřijovskỳ J, Hartvich F (2017) Uav-based optical granulometry as tool for detecting changes in structure of flood depositions. Remote Sensing 9(3):240

Murtiyoso A, Grussenmeyer P, Landes T, Macher H (2021) First assessments into the use of commercial-grade solid state lidar for low cost heritage documentation. In: XXIV ISPRS Congress (2021 Edition), 5–9 Juillet 2021, Nice (en Ligne), vol 43

Purinton B, Bookhagen B (2019) Introducing pebblecounts: a grain-sizing tool for photo surveys of dynamic gravel-bed rivers. Earth Surf Dyn 7(3):859–877

Remondino F, El-Hakim S (2006) Image-based 3d modelling: a review. Photogram Rec 21(115):269–291

Teppati Losè L, Spreafico A, Chiabrando F, Giulio Tonolo F (2022) Apple lidar sensor for 3d surveying: tests and results in the cultural heritage domain. Remote Sensing 14(17):4157

Vincent L, Soille P (1991) Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans Pattern Anal Mach Intell 13(06):583–598

Wang D, Watkins C, Xie H (2020) Mems mirrors for lidar: A review. Micromachines 11(5):456

Westoby MJ, Dunning SA, Woodward J, Hein AS, Marrero SM, Winter K, Sugden DE (2015) Sedimentological characterization of antarctic moraines using uavs and structure-from-motion photogrammetry. J Glaciol 61(230):1088–1102

Zhao L, Zhang S, Huang D, Wang X, Zhang Y (2020) 3d shape quantification and random packing simulation of rock aggregates using photogrammetry-based reconstruction and discrete element method. Constr Build Mater 262:119986