Extraction of key geometric parameters from segmented masonry arch bridge point clouds

Yixiong Jing, Brian Sheil, Sinan Acikgoz

Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ,
(yixiong.jing@wolfson.ox.ac.uk; brian.sheil@eng.ox.ac.uk; sinan.acikgoz@eng.ox.ac.uk)

Key words: p*oint cloud; segmentation; deep learning (DL); BridgeNet; synthetic dataset; feature extraction*

ABSTRACT

Masonry arch bridges constitute the majority of the European bridge stock. Most of these bridges were constructed in the 19th century and feature a wide range of geometric characteristics. Since construction drawings rarely exist, the first step in the assessment of these bridges is the characterisation of their in-situ geometry, which may involve significant geometric distortions. In recent years, LIDAR devices have been widely used by bridge owners due to their ability to remotely and rapidly collect point cloud data. To enable the engineering assessment practice to benefit from this data, this research uses the recently developed deep learning (DL) neural network BridgeNet to autonomously segment masonry bridge point clouds into different components. Due to the limited availability of 3D point clouds, BridgeNet is trained using a synthetic multi-span masonry arch bridge dataset; the network is then tested on real arch bridge point clouds. By fitting appropriate primitive shapes to bridge component point clouds using Random Consensus Sampling (RANSAC) techniques the bridge geometry is effectively characterised by a few parameters.

I. INTRODUCTION

In the UK, there are approximately 18,000 operational masonry railway bridges which occupy 47% of the total bridge stock (Orbán, 2004). Increasing traffic loading has been hastening the rate of deterioration of these bridges. Therefore, regular inspections are required to ensure their safety (Acikgoz *et al*., 2018). The evaluation of loading capacity requires detailed knowledge of the geometry of the structure, such as the radius of curvature of the arch. However, construction drawings are often unavailable for masonry arch bridges and measurements of geometric characteristics are sometimes limited by access issues.

Laser scanning provides an efficient and non-contact method of obtaining the in-situ 3D geometry of large-scale civil engineering infrastructure and representing them by using point clouds. Key geometric parameters describing the bridges can then be estimated by post-processing the point clouds (Riveiro *et al.*, 2011; Schnabel *et al*., 2007). However, these estimations often require the manual segmentation of the point clouds which is time-consuming for large datasets. Therefore, an automatic algorithm is necessary for conducting both semantic and instance segmentation. For instance, to determine the radius of curvature of an arch, points belonging to the arches need to be first semantically classified and then clustered into different arch instances. Traditional feature-based segmentation algorithms usually make assumptions based on the shape and topological relationships of each component (Lu *et al*., 2019; Riveiro *et al*., 2016). Although these algorithms can achieve good segmentation accuracy, they cannot be generalized to other datasets since they are designed for specific geometric characteristics.

Compared with traditional segmentation algorithms, deep learning (DL) algorithms are known for their ability to offer generalized solutions that are robust to noise. One of the most popular 3D neural networks, PointNet++ (Qi *et al*., 2017), achieves good performance on semantic segmentation with the multi-scale feature extraction design. However, the structure of PointNet++ only allows it to process less than ten thousand points at the same time while a common masonry arch bridge point cloud may contain millions of points. Recently, RandLA-Net (Hu *et al*., 2020) was proposed to conduct semantic segmentation on large-scale point clouds efficiently by using random point sampling and a lightweight feature extractor. Similarly, FG-Net (Liu *et al*., 2020) uses sophisticated yet efficient feature extractors to group local and global features which achieve the best results to date on the public partial segmentation dataset called PartNet (Mo *et al*., 2019). Based on their work, BridgeNet (Jing *et al*., 2022) was developed to perform segmentation on large-scale masonry arch bridge point clouds. To tackle data scarcity, a synthetic point cloud simulator was developed in the same study and it was used to generate a large synthetic training dataset. The trained algorithm was tested on real point clouds of seven masonry arch bridges from the UK. Semantic segmentation results were post-processed with a simple and memory-efficient clustering method called DBSCAN (Ester *et al*., 1996) to perform instance segmentation.

To demonstrate how point cloud segmentation results can be used in the assessment of masonry arch bridges, the automated extraction of key geometric parameters from segmented point clouds is

investigated in this study. More specifically, key geometric parameters are extracted from BridgeNet segmented point clouds (Jing *et al*., 2022). For the arch point clouds, the radius and direction of the generatrix are estimated by idealizing arch intrados geometry with partial cylinders. Some shape fitting algorithms rely on principal component analysis to identify the direction of the generatrix (Nurunnabi *et al*., 2017; Proença *et al*., 2018). However, these algorithms rely on the assumption that the length of the cylinder is larger than the radius of curvature, which is unsuitable for many masonry arch bridges. Instead, a RANSAC-based algorithm is used in this study by iteratively fitting a small group of points with a cylinder function and excluding outliers to obtain the best-fit geometry. Piers are combinations of multiple planes that sometimes feature complex geometric characteristics such as skewed flank surfaces. Another RANSAC-based algorithm is used to identify planes from pier instances and the skew angles of flank surfaces are then extracted from the planes.

The paper is organized as follows. Section II introduces the BridgeNet used for performing segmentation on masonry point clouds and the real and synthetic datasets. In Section III, training results from BridgeNet are presented. Section IV presents the extraction of key geometric parameters of arches and piers. Section V provides the conclusion of this paper.

II. BRIDGENET AND DATASET

A. *BridgeNet*

BridgeNet adapted the architecture from RandLA-Net (Hu *et al*., 2020). RandLA-Net (Hu *et al*., 2020) was proposed to perform semantic segmentation on large-scale point clouds efficiently. Instead of applying time-consuming sampling methods such as the farthest point sampling, it uses random point sampling to downsample the point clouds. To compensate for the lost information, a lightweight feature extractor was introduced to explicitly concatenate geometric information with encoded features. However, it is observed in our experiments that the feature extractor in RandLA-Net (Hu *et al*., 2020), whilst achieving good accuracy in semantic segmentation such as city scene, could not satisfactorily capture the level of local feature details required in our segmentation tasks. Therefore, a more sophisticated feature extractor module was required. To obtain an optimal balance between computational efficiency and performance, we adopted the feature downsampling method from FG-Net (Liu *et al*., 2020). As shown in Figure 1, the input has the size of $N_i \times C_{in}$ in which N_i represents the number of points and C_{in} is the input feature dimension. To achieve memory efficiency, the feature dimension is first downsampled by a ratio of *m* (8 in our case) using a convolutional layer. Ball query is then used to group K neighbours of each point to capture the local relationship. The grouped feature tensor is finally

passed to convolutional layers to encode the local features and the output has the size of $N_i \times K \times C_{out}$, in which C_{out} is the output feature dimension. It is notable that step 3 in Figure 1 consumes the most GPU memory. Therefore, the downsampling of feature dimension alleviates the memory consumption by condensing the feature tensor into 1/8 of its original size. More details can be found in FG-Net (Liu *et al*., 2020).
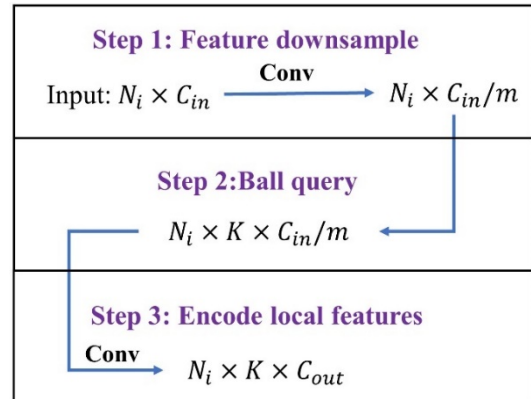

Figure 1. Illustrations of the feature extractor to encode local features of large-scale point clouds; the input is first downsampled in feature dimension by the ratio of *m*, *K* neighbour points are then grouped for each point; the clustered features are finally passed through a convolutional layer for the output.

The input size of our neural network is $N \times 3$, in which N is the number of points and 3 are the raw *x, y* and *z* point coordinates. The input is first passed through an encoder which is composed of four feature extractor layers to learn local features in different resolutions. The random point sampling adopted from RandLA-Net (Hu *et al*., 2020) is applied after every single layer to downsample point clouds with a ratio of 4, which achieves time efficiency for large-scale point clouds. The features are then sent to a decoder to scale the number of points back to the original input by using the nearest interpolation method (Qi *et al*., 2017). A 1×1 convolution layer is subsequently applied to reshape the feature dimension corresponding to the semantic label with the size of $N \times C_{sem}$, in which C_{sem} represents the number of semantic classes. Details of BridgeNet can be found in (Jing *et al*., 2022).

B. *Dataset*

The dataset is composed of synthetic and real components, which are used for training and testing respectively.

The synthetic dataset is designed to represent the geometry of real masonry arch bridges. The geometry of each component for the masonry arch bridge is parameterized to automatically generate thousands of point clouds with different shapes. The noiseless synthetic point cloud produced in this way is then corrupted to simulate construction errors and geometric irregularities. For example, the height of the

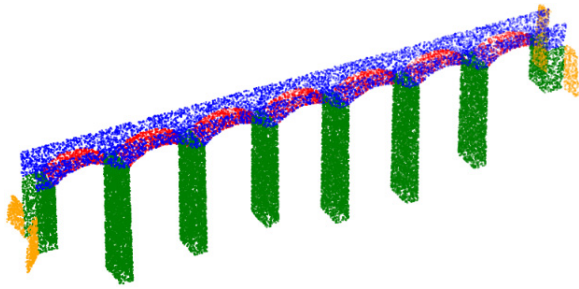pier is randomized to represent variations caused by irregular topography as shown in Figure 2.

Figure 2. Example of synthetic masonry arch bridge point cloud with randomized geometry featuring different pier heights, span lengths and wingwall geometries.

The real dataset contains point clouds of seven multi-span masonry arch railway bridges from the UK. Marsh Lane Viaduct is shown in Figure 3 as an example in which the vegetations, parapet and poorly scanned areas are removed since they are not considered in the synthetic data simulator. The processed point clouds are then classified into five components which are the spandrel wall, pier/abutment, arch, hole and wingwall. The manually segmented Marsh Lane Viaduct point cloud data in Figure 4 features the first four components, as the viaduct did not have wingwalls. Other bridges in the real dataset include this feature.

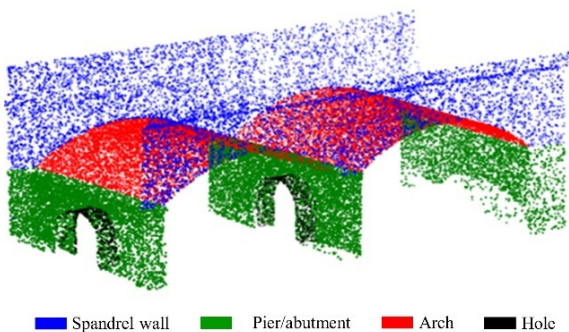Figure 3. Example of the real point cloud of Marsh Lane Viaduct used in the test dataset.

| ■ Spandrel wall | ■ Pier/abutment | ■ Arch | ■ Hole |

Figure 4. Manual semantic segmentation of the point cloud of Marsh Lane Viaduct into four primary components.

III. TRAINING EXPERIMENTS

A. *Implementation details and metrics*

The training dataset contains only synthetic data sampled from the synthetic data generator introduced in Section II. BridgeNet is trained with an increasing number of synthetic bridge point clouds from 500 to 3000. The trained BridgeNet is then tested on all seven real bridges each of which contains millions of points. The network can efficiently segment each point cloud in seconds. The experiments are conducted on an RTX 3080 GPU with 10 GB of memory.

Mean Intersection-over-Union ($mIoU$) is adopted in the experiments to assess the performance of the network as shown in Equation 1, in which TP_i is the true positive for class i, FP_i is the number of points being misclassified as class i, FN_i is the number of misclassified points belonging to class i and n_{cls} is the number of components for the corresponding point cloud.

$$mIoU = \frac{\sum_{i=1}^{n_{cls}} \frac{TP_i}{TP_i + FP_i + FN_i}}{n_{cls}} \qquad (1)$$

$mIoU$ can identify poorly segmented classes even when they occupy only a very small portion of the total point cloud such as the hole or wingwall classes.

B. *Semantic segmentation results*

The influence of an increasing number of point clouds on the segmentation accuracy of BridgeNet is shown using the $mIoU$ results in Table 1. The best accuracy of 0.779 is achieved using 2000 point clouds during training. As the size of the training dataset is increased, the $mIoU$ is first improved which indicates that the geometric information provided by the synthetic data provides useful training information. However, the $mIoU$ for the case with 3000 training point clouds demonstrates poorer performance and experience overfitting related problems. Hence, the network trained with 2000 training point clouds will be used in the rest of the paper.

Table 1. $mIoU$ of BridgeNet on test dataset with increasing training samples in the synthetic dataset

	$mIoU$
500	0.654
1000	0.682
1500	0.751
2000	0.779
3000	0.737

By way of example, the segmentation results for Marsh Lane Viaduct are visualized in Figure 5. The true labels are presented first, then the BridgeNet segmentation results. The hole in the Marsh Lane viaduct is mostly misclassified as a pier element. This could be attributed to the unbalanced proportions of classes in the dataset. The network tends to classify points belonging to holes into piers since the points of holes only occupy a small portion of the whole point cloud. However, the general shape of each component

is well captured for all data in the test set. The next section evaluates the accuracy of geometric parameters extraction in the presence of segmentation inaccuracy.
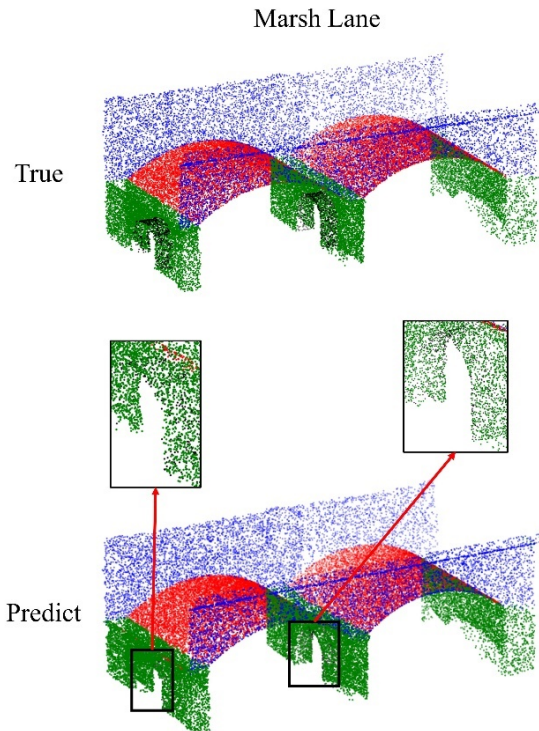
Marsh Lane

Figure 5. Visualizations of the true segmentation label of the Marsh Lane Viaduct compared with the semantic segmentation results from BridgeNet. Misclassifications are highlighted and shown inset with a zoomed-in view within black bounding boxes.

IV. GEOMETRIC EXTRACTION

A. Instance segmentation

Since the network only gives the semantic segmentation results, the unsupervised learning package DBSCAN (Pedregosa et al., 2011) is used to perform instance segmentation to cluster points into separate objects as shown in Figure 6. The outliers were automatically removed by filtering out clusters that contain points lower than the defined threshold. The hyper-parameters used in the algorithm had to be adjusted following a trial and error approach to achieve the best instance segmentation performance.

B. Geometric feature extraction of arches

Geometric parameters are extracted from the 'ground truth' point cloud of arches, which were manually segmented. These are compared with the geometric parameters obtained from the arch point clouds segmented by BridgeNet and DBSCAN to demonstrate the feasibility of the proposed framework. Since the arch is assumed to approximate a partial cylinder, the radius and direction of the generatrix are extracted from the point clouds for comparison.
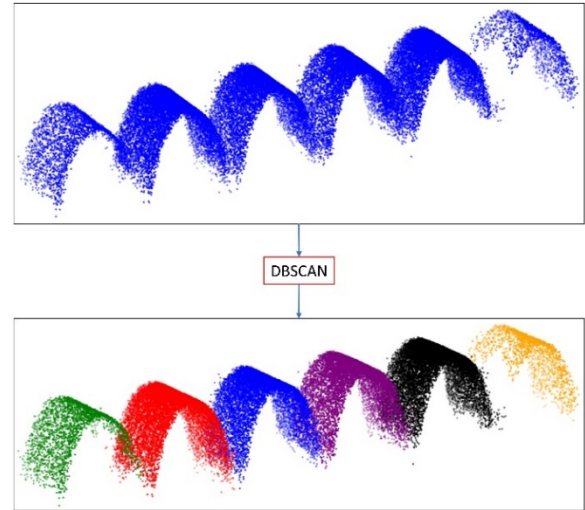

Figure 6. Instance segmentation performed after semantic segmentation for arches of the Chelmsford Viaduct by using DBSCAN.

Algorithm 1. Robust extraction of radius and generatrix of arches based on RANSAC and a cylinder fitting function

---

Input: point cloud for a single arch instance
$P_{arch} = \{p_i\}, i = 1, 2, \ldots, N, p_i = (x_i, y_i, z_i)$, where N is the number of points; maximum iteration n ; maximum radius error tolerance e_{radius_tol} ; minimum number of points N_{cy_tol} satisfying the radius difference tolerance r_{tol} compared with the radius from the fitted cylinder function; several points chosen from P_{arch} to fit the cylinder function N_{cy_tol}

Output: radius r_{out} and generatrix direction d_{out} , in which $d_{out} = (x_{out}, y_{out}, z_{out})$

$m = 0$
while $m < n$ **do**
  Randomly sampling N_{cy} points from P_{arch} which produces a new point set of P_{cy};
  r_i, d_i, c_i = cylinder_fitting(P_{cy}), which c_i is the centre of the middle cross-section of the cylinder in iteration i;
  Calculate the mean absolute deviation of the radius e_{radius} between the rest of the points in P_{arch} by using the fitted r_i, d_i, c_i;
  Calculate the number of points N_{cy} fitted to the cylinder which satisfies the radius tolerance r_{tol};
  if $N_{cy} > N_{cy\_tol}$ **and** $e_{radius} < e_{radius\_tol}$ **then**
    $r_{out} = r_i, d_{out} = d_i$;
    return r_{out}, d_{out};
  else then
    m = m + 1;
    Continue

The two parameters can be calculated by using a RANSAC-based iterative geometric extraction algorithm as shown in Algorithm 1. By randomly selecting a small set of points (15 points in our case) from the arch point cloud, the algorithm first fits the 3D parametric cylinder function for the radius r_i, generatrix direction d_i and the centre coordinates c_i in the middle cross-section. For the rest of the points, the shortest distance between them and the fitted cylinder generatrix (uniquely determined by c_i and d_i) is then calculated.

By subtracting r_i from this value, the residual error of the radius is obtained for each point. The absolute value of errors is summed and averaged to obtain the mean absolute deviation e_{radius}. The number of points on the cylinder, N_{cy}, can then be determined by filtering points with residual errors larger than a pre-defined threshold r_{tol}. If N_{cy} is larger than N_{cy_tol} and e_{radius} is smaller than e_{radius_tol}, the fitted r_i and d_i are accepted and the loop ends. Otherwise, the loop continues until the maximum number n is reached.

The extracted geometric parameters in the final step of the algorithm for the segmented arches are compared to the parameters extracted from ground truth arches as shown in Table 2. Root mean squared errors (RMSE) are specified since the geometric extraction was conducted for multiple arches. The RMSE for the radius is typically smaller than 3 cm, indicating remarkable accuracy. The direction of the generatrix is captured with an RMSE of < 0.2° which highlights the robustness of the segmentation results.

Table 2. RMSE values associated with the radius and the generatrix angle obtained by fitting cylinders to arch point clouds segmented using BridgeNet are presented. The errors are relative to the radius and generatrix estimations from cylinders fitted to manually segmented ground truth point clouds

	Radius [m]	Generatrix [°]
Chelmsford	$8.43*10^{-3}$	$1.70*10^{-1}$
Hertford	$6.08*10^{-3}$	$1.57*10^{-1}$
Peter 0	$1.87*10^{-2}$	$8.57*10^{-2}$
Peter 1	$1.30*10^{-2}$	$1.70*10^{-1}$
Digswell	$2.79*10^{-2}$	$1.87*10^{-1}$
Stapleton	$1.91*10^{-2}$	$9.29*10^{-2}$
Marsh Lane	$2.62*10^{-2}$	$4.61*10^{-2}$

C. *Geometric feature extraction of piers*

Skewed pier flank surfaces are a feature of the Digswell Viaduct in the real dataset as shown in Figure 7. Therefore, this section considers Digswell Viaduct as an example to demonstrate the robustness of the proposed extraction process to capture complex geometric features. The skew angle is defined as the angle between the vector normal to the pier flank surface and the horizontal plane. Skew angles are first extracted from manually segmented planes as the ground truth values for subsequent comparison with predicted values.

Another RANSAC-based algorithm is used to separate planes of segmented pier point clouds. Three points are randomly chosen from point clouds to fit a 3D surface and the rest of the points would be checked whether they are on the plane or not based on a distance threshold. After reaching the maximum number of iterations or sufficient points being fitted to the 3D surface, it is accepted as the current best plane and all points belonging to this plane are removed from the pier point cloud. By repeating this procedure until either the maximum number of planes is reached (four, for typical piers) or not enough points are left in the point cloud, the algorithm then stops and returns all extracted planes. An example of the plane extraction is visualized in Figure 8 where four planes are represented by red, grey, orange and black points.
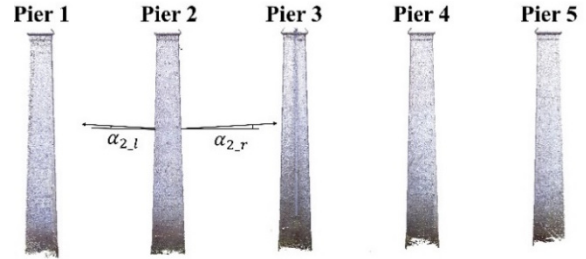

Figure 7. Piers of Digswell Viaduct, in which the flank surfaces are skewed. There are a total of five scanned piers and the right and left skew angles of pier 2 are denoted by α_{2_r} and α_{2_l} respectively.

Figure 8. Extractions of planes by using RANSAC-based algorithm for pier 2, four planes are represented by red, grey, orange and black points.

The flank surfaces along the bridge transverse direction are selected for calculating the angle of skew since they contained more points. As shown in Figure 7, the left and right skew angle for pier i is represented by α_{i_l} and α_{i_r}. Table 3 demonstrates that the skew angles of flanks are usually smaller than 2° and the total error is represented by the mean average deviation (MAD) of 0.06° which demonstrates the excellent accuracy of the proposed framework.

V. CONCLUSION

In this paper, we proposed an automated and generalisable framework for extracting geometric features from masonry arch bridge point clouds to facilitate the assessment and inspection. Since there is no available or published neural network to tackle with partial segmentation of large-scale point clouds, BridgeNet (Jing *et al.*, 2022) was used to provide a general solution for infrastructure segmentations. Due to the scarcity of real data for training, that work developed a masonry arch bridge point cloud generator to create synthetic point clouds with different geometries. The synthetic point clouds were used to train the network, which was later tested on seven real railway bridge point clouds. The resulting segmentation

2022, Editorial Universitat Politècnica de València

had good accuracy but featured some errors, which could influence geometric parameter extraction.

Table 3. The extracted skew angles by fitting the plane function with the manually segmented ground true pier point clouds, predicted ones and the absolute errors between them

		True [°]	Predict [°]	Abs errors [°]
Pier 1	α_{1_l}	--	--	--
	α_{1_r}	0.83	0.87	0.04
Pier 2	α_{2_l}	1.98	2.03	0.05
	α_{2_r}	1.05	0.99	0.06
Pier 3	α_{3_l}	1.89	1.93	0.04
	α_{3_r}	0.93	0.93	0.00
Pier 4	α_{4_l}	1.87	1.82	0.05
	α_{4_r}	1.04	0.93	0.11
Pier 5	α_{5_l}	1.85	1.95	0.10
	α_{5_r}	--	--	--
MAD	0.06			

To investigate the feasibility of using the BridgeNet segmentation results for extracting key geometric parameters of bridges, both true and predicted arches are fitted with partial cylinders using a RANSAC-based approach. Another RANSAC-based algorithm is used to separate and select flank surfaces of predicted piers to obtain the skew angle. The geometric parameters extracted from the segmentation results of BridgeNet were compared to true values, demonstrating high accuracy. The results emphasize the robustness of segmentation and highlighted the useful geometric information that can be gained from segmented point clouds.

VI. ACKNOWLEDGEMENTS

References

Acikgoz, S., DeJong, M. J., and Soga, K (2018). Sensing dynamic displacements in masonry rail bridges using 2D digital image correlation. Structural Control and Health Monitoring, Vol. 25, No 8, pp. e2187.

Ester, M., Kriegel, H. P., Sander, J., and Xu, X (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining (KDD'96)*, AAAI Press, Vol. 96, No. 34, pp. 226-231.

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni N and Markham, A (2020). Randla-net: Efficient semantic segmentation of large-scale point clouds. In: *Proc. Of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11108-11117.

Jing, Y., Sheil, B.B., and Acikgoz, S (2022). Segmentation of large-scale masonry arch bridge point clouds with a synthetic bridge simulator and the BridgeNet neural network. Under review.

Liu, K., Gao, Z., Lin, F., and Chen, B. M (2020). FG-Net: Fast Large-Scale LiDAR Point Clouds Understanding Network Leveraging Correlated Feature Mining and Geometric-Aware Modelling. arXiv preprint.

Lu, R., Brilakis, I., and Middleton, C. R (2019). Detection of Structural Components in Point Clouds of Existing RC Bridges. *Computer-Aided Civil and Infrastructure Engineering*, Vol. 34, No. 3, pp. 191-212.

Mo, K., Zhu, S., Chang, A. X., Yi, L., Tripathi, S., Guibas, L. J., and Su, H (2019). PartNet: A Large-scale Benchmark for Fine-grained and Hierarchical Part-level 3D Object Understanding. In: *Proc. Of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 909-918.

Nurunnabi, A., Sadahiro, Y., and Lindenbergh, R (2017). Robust cylinder fitting in three-dimensional point cloud data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-1/W1, pp. 63-70.

Orbán, Z (2004). Assessment, reliability and maintenance of masonry arch railway bridges in Europe. *Arch Bridges IV– Advances in Assessment, Structural Design and Construction*. Eds: P. Roca and C. Molins, Barcelona, pp. 152-161.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., and Duchesnay, E (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, Vol. 12, pp. 2825-2830.

Proença, P. F., and Gao, Y (2018). Fast cylinder and plane extraction from depth cameras for visual odometry. In: *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 6813–6820.

Qi, C. R., Yi, L., Su, H., and Guibas, L. J (2017). Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In: *Proc. Adv. Neural Inf. Process. Syst.*, pp. 5099–5108.

Riveiro, B., DeJong, M. J., and Conde, B (2016). Automated processing of large point clouds for structural health monitoring of masonry arch bridges. *Automation in Construction*, Vol. 72, pp. 258-268.

Riveiro, B., Morer, P., Arias, P., and De Arteaga, I (2011). Terrestrial laser scanning and limit analysis of masonry arch bridges. *Construction and Building Materials*, Vol. 25, No. 4, pp. 1726-1735.

Schnabel, R., Wahl, R., and Klein, R (2007). Efficient RANSAC for Point-Cloud Shape Detection. *Computer Graphics Forum*, Vol. 26, No. 2, pp. 214-226.