euspen's 23rd International Conference & Exhibition, Copenhagen, DK, June 2023
www.euspen.eu

An intelligent surface segmentation approach based on U-Net for structured and freeform surface characterisation

Weixin Cui [1], Wenhan Zeng [1], Shan Lou[1], Paul J. Scott[1], Xiangqian Jiang[1]

[1]*EPSRC Future Metrology Hub, Centre for Precision Technologies, School of Computing and Engineering, University of Huddersfield, Huddersfield, HD1 3DH, UK*

Weixin.Cui@hud.ac.uk

Abstract

With the freeform structured surface becoming more dominant in the engineering surface metrology and manufacturing as they have deterministic patterns which designed with some specific functionalities to meet the engineering requirements such as optical, electric contact and bearing properties, the ability to adequately characterise them is crucial for optimising the performance through reducing cost and achieving precise control of such specifically functional components. A general surface characterisation scheme for complex freeform includes three operations: form removal, denoising, segmentation. Following the first two steps, the surface measurement will be converted to a scale-limited surface. The next crucial step is segmentation, which separates the surface topography into a number of non-intersecting regions so that they can be analysed separately and in relation to one another, for example, by computing shape attributes or its pertinent dimensions. Traditional computer-vision methods such as watershed and active contour approaches perform well on this task, but these algorithms have high computational complexity with long running time and require test phase for fine-tuning and rely on users-judgement. Additionally, inappropriate initial conditions will lead to over- or under- segmentation. Deep learning-based segmentation techniques become dominated with outstanding performance and higher accuracy. Therefore, we proposed a novel deep learning-based surface segmentation method for the freeform structured surfaces which based on the U-Net model and data augmentation techniques are utilised to enlarge the raw dataset of twenty surface measurements. The training data include converted RGB-images of surfaces and corresponding feature masks which is the ground truth pixel labels generated using computer vision techniques including thresholding and edge operators. Once the model has been trained, it can output the drawn feature map of input surface with precise boundaries efficiently. The U-Net segmentation model is a kind of Encoder-decoder architecture with benefits of lower cost and high efficiency with great segmentation accuracy using few training data. The experimental results show the remarkable and applicable performance of structured surface segmentation to meet the metrological requirements, which can support the intelligent surface characterisation framework and for further feature attributes analysis and parameterisation.

Keywords: Surface segmentation, freeform structured surface, deep learning, U-Net, surface characterisation

1. Introduction

In the era of Industry 4.0, intelligent manufacturing is rapidly evolving by leveraging real-time data analysis and artificial intelligence (AI) techniques to optimise the manufacturing process. As a sub-scope of manufacturing, metrology, the science of measurement, is also advancing towards digitalization and intelligence. The future development of metrology faces two key challenges: integrating metrology into the design process [1] and establishing smart data analytics systems [2]. In the next few decades, metrology is expected to improve the understanding of complex systems through digital surface texture analysis technologies and support decision-making frameworks using AI-assisted techniques. The former requires establishing connections between surface features and functionality, while the latter is anticipated to lead to the application of machine learning to new areas. Surface characterisation is a critical component of data analytical systems in metrology, allowing the optimisation of the functional performance of components and reducing costs. Engineering surfaces are transitioning from stochastic to deterministic patterns, referred to as freeform structured surfaces, with complex forms and structured features designed to meet specific functions such as good bearing properties, electrical contact, and optical properties [3]. The general surface characterization scheme for freeform structured surfaces comprises three main operations: form removal, denoising, and segmentation, as shown in Figure 1.

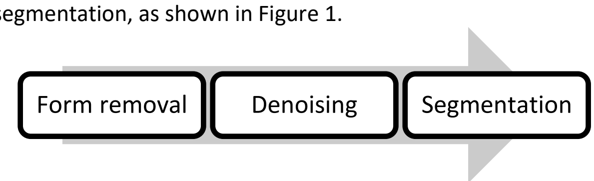

Figure 1. Surface charaterisation sheme

The initial step is form removal to remove reference form as large-scale components from primary surface measuments. Next, denoising targets small-scale components and suppresses high-frequency noise to obtain a scale-limited surface for feaure extraction. Segmentation is then performed on the scale-limited surface to divide the surface topography into a number of non-intersecting regions and use a set of parameters to identify the individual feature attributes or relations between them [4]. The future development of charaterisation is integrating into the smart metrology system to achieve real-time data analytics and optimise manufacturing process which means requiring design an intelligent segmenter to process different type of surfaces measurements.

Traditional segmentation methods are usually based on computer vison techniques such as edge operators, active contour, watershed and level sets. While they perform well on different types of sufaces, but they usually involve manual operations for each individual input such as initial condition setting or parameter tuning. Inproper initial conditions will lead to over- or under- segmentation which implies techniqual background is required for users to apply these algorithms. In addition, they are inefficienct for multiple imags application due to computational complexity. Watershed segmentation is a commonly used method used in surface metrology [5]. Take it as an example, there are two ways: top-down and bottom-up depending on different tasks and it requires the selection of user-defined markers for segmentation. The main drawback of watershed and active contours is over-segmentation due to the local minimas sensitivity [6].

To address these issues, the ideal feature extractor should achieve automatic segmentation and process multiple inputs simultaneously for diverse surfaces without any manual operation. Deep learning techniques have been successfully applied in image processing field including segmentation task, with benifits of intelligence and generalisation. In this study, we apply the U-Net model, a well-known network for image semantic segmentation, conbining data angmentation and transfer learning techniques to perform on freeform structured surface using only a small amount of dataset. The experimental results indicate that retrained UNet model can reach high accuracy of segmentation and classify structured features simultaneously. No need for manual setting and tuning for model use, which implies it can function as an intelligent segmenter to achieve automtic segmentation for multiple inputs. Due to the GPU support for deep learning, compared to the traditional approachs, the proposed learning-based method can output results efficiently once the model has been trained. Therefore, it is also friendly to non-expert users to apply.

2. Brief literature review of image segmentation

This section will give a brief literature review of traditional used surface segmentation methods and advanced deep learning-based models used for image segmentation. Commonly used edge operators including Sobel, Roberts and Prewitt are useful to detect feature edges by quantifying gradient vector of each pixel that represent local changes in pixel-values intensity and they usually used before segmentation as a pre-step to enhance accuracy and performance [7]. Active contour models (ACMs) have shown better performance as represented by the active contour without edge model [8]. Furthermore, Chan and Vese introduced a new approach using level set functions to formulate the segmentation model treated as an energy minimisation objective which can be solved through solving PDEs [9]. Watershed transform algorithm is a robust algorithm based on mathematical morphology as with adaptive thresholding, from a threshold is too slow for dividing objects and gradually raised to an optimal level. Scott proposed an extended watershed method based on Maxwell's theory and Pfaltz graph, which expands input range from areal height maps to true 3D mesh surfaces [5]. Construction is an operation to merge over-segmented regions by Wolf pruning.

Currently, there are only a limited number of applications for machine learning techniques in surface metrology, however, due to the similarities in data structure, deep learning methods have high potential in surface metrology as their poweful capability in performing image processing tasks. From the view of tasks in deep learning, image segmentaion can be formulated as the problem of classifying elements with semantic labels (semantic segmentation) or partitioning of individual objects (instance segmentation) or both (panoptic segmentation). For structured engineering surfaces, patterns are structured and features need to be identified individually to match various functionalities such as optical and electrical property. Hence, applying semantic segmentation can not only identify boundary of structures but also gather information with features relationships. One of the first successful deep learning-based approaches for image segmentation was the Fully Convolutional Network (FCN) proposed by Long et al. in 2015 [10]. FCN uses a convolutional neural network (CNN) that is trained end-to-end to predict the segmentation mask of an input image. FCN was able to achieve state-of-the-art results on several benchmark datasets, demonstrating the potential of deep learning for image segmentation. Following the success of FCN, several other deep learning-based approaches for image segmentation were proposed. One most popular architecture is U-Net, proposed by Ronneberger et al. in 2015 for biomedical images segmentation application [11], which uses Encoder-decoder structure and has been shown to perform well on small training dataset. Mask R-CNN is another approach proposed by He et al. in 2017 [12] as an extension of the Faster R-CNN of which is developed for instant segmentation and has achieved state-of-the-art results on several benchmark datasets. Additionally, DeepLab [13] is another family of models that use atrous convolution and a multi-scale feature fusion approach to combine features from different levels of the network. The V-Net is another classic model based on FCN model proposed by Milletari et al. [14] which is used for 3D image segmentation. Some hybrid methods are combining CNNs with classic computer vision approaches such as watershed or active contour [15] to perform different tasks. In this metrological application, the U-Net model was selected because it offers two key advantages. Firstly, it has the ability to attain high accuracy in pixel-level segmentation while still maintaining important contextual and location information. Secondly, it performs well even when trained on only a limited number of samples, which is ideal for this particular application as there is a scarcity of available engineering surface data.

3. Methodology

3.1. UNet architecture

Originally employed for medical image segmentation, U-Net is a traditional neural network that has proven successful in a number of semantic segmentation applications. As seen in figure 2, the fundamental architecture is a two-path encoder-decoder. By encoding feature channels in higher resolution layers, the encoder component is a path that contracts to capture the context information of inputs. To enable accurate localization from feature representations, the decoder component is a similar expanding path connected encoder blocks symmetrically. It should be noted that the skip connections technique is employed here to combine the output of the same-level decoder layer with the feature maps of each encoder layer in order to increase the localization accuracy. Consisting of two 33 convolution kernals, a rectified linear unit (ReLU), and a 22 max pooling for downsampling but doubling the feature channels, contraction is a repetitive structure. Expansion is achieved by regular convolutions and upsampling to same dimensionality using transposed convolutions. For surface segmentation application, the inputs of model are RGB images $(r \times c \times 3)$ with ground truth feature masks $(r \times c \times 1)$, and the output are pixel-labeled category maps. The Tversky loss $L = 1 - TI$ [16] based on Tversky Index (TI) is used here to measure the overlap between predictions and labels.

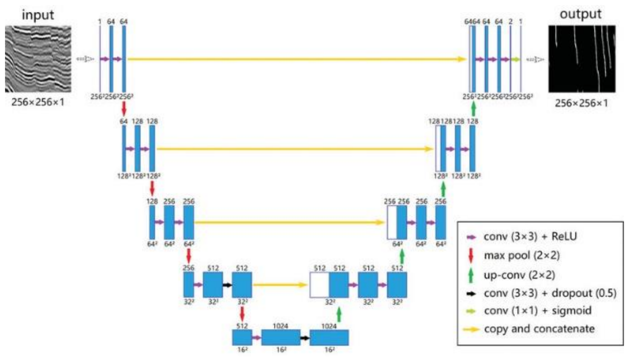

Figure 2. Original UNet architecture [11]

3.2. Proposed method framework

An intelligent surface segmenter is proposed based on U-Net model with utilisation of data augmentation and transfer learning techniques. Data augmentation is used to expand original dataset and resize to uniform the inputs. Transfer learning is employed here based on pre-trained network to speed up the training and reduce the cost. The method framework is shown in figure 3 and implementation steps are summarised as follows:

1) Create image datastore with original input RGB images converted from surface measurements.
2) Generate ground truth labels using the image processing techniques for each surface.
3) Use data augmentation techniques including cropping and warping transformations to obtain augmented training dataset containing patches and corresponding pixel-labeled feature masks.
4) Load pre-trained UNet network and replace the final layers with customised feature classes and image size.
5) Retrain the model based on our dataset and oprimise the hyper-parameters to obtain best-tuned model.
6) Test our model based on test dataset and evaluate by segmentation metrics.

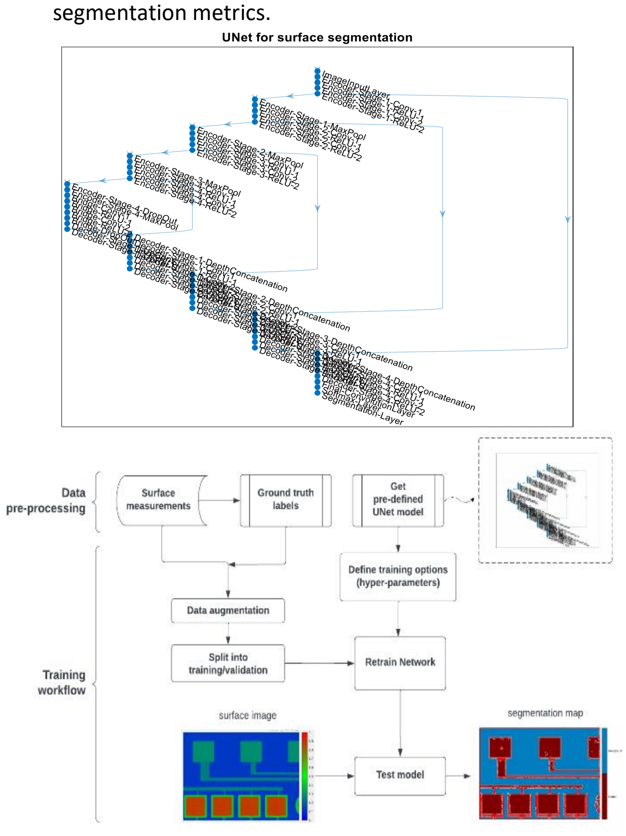

Figure 3. Framework of proposed UNet-based surface segmentation method

4. Experiments

4.1. Surface segmentation

The U-Net model is constructed and trained in MATLAB (2022b) platform. As a result of the limited engineering samples of structured surfaces, the original data contains thirty-seven structured and freeform surface measurements featured various types of structures. Then split into training, validation and test dataset in a ratio of 60%:20%:20%. In order to expand dataset, data augmentation including cropping and rotating is carried out, resulting in expansion of each sample into 50 patches. This leads to training dataset being expanded to 2000 patches of size $128 \times 128 \times 3$, along with their corresponding pixel-labeled masks of size $128 \times 128 \times 1$ as figure 4 shown. The validation and test dataset contain 400 patch-pairs of the same size.

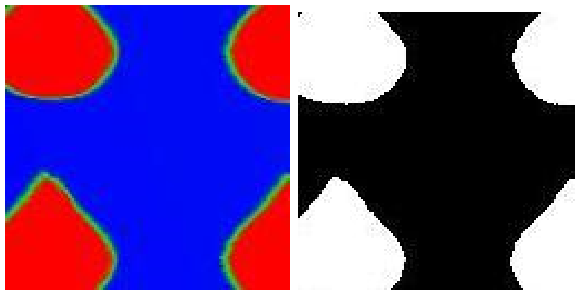
Figure 4. Training patch with label mask

The experimental results are visualised with boundary trace. Figure 5 shows typical cases from MEMS surface. The segmentation results indicate that our model can extract the feature boundary accurately especially for overlapped features and output precise pixel-level location information. Figure 6 shows a type of freeform structured surface named 'Simon-star'. As we can see, it has distortion points due to presence of noise and which also demonstrates the lack of robustness. Based on model output it is can calculate the field and feature parameters of surface based on location information with resolutions.

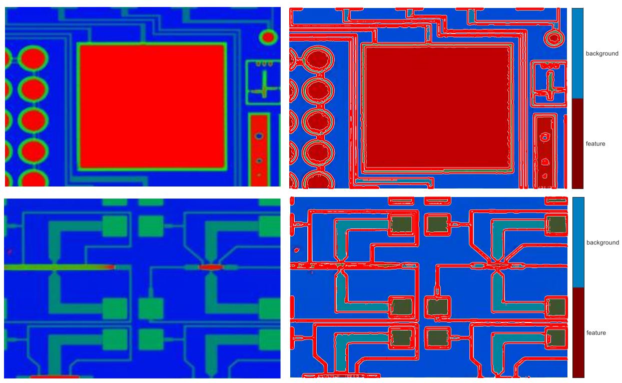
Figure 5. U-Net segmentation performance for MEMS surfaces

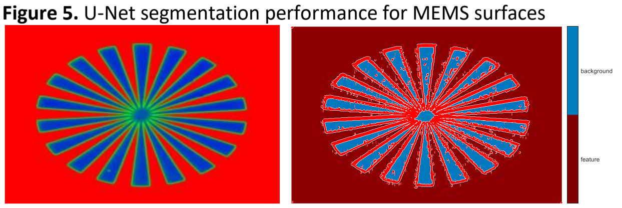
Figure 6. U-Net segmentation performance for 'Simon stars' surface

Therefore, the results are compatible with next step of paraterisation. For each input, we can obtain segmentation result within average one second on CPU and a hundred of milliseconds on GPU. The specific running time depends on input size. No manual involved operation which indicates the model can support processing multiple inputs efficiently.

4.2. Evaluation

To quantitively assess the segmentation performance on surface metrology, three metrics are calculated to instuitively measure the quality and accuracy of our model. Commonly used metrics for semantic segmentation task are Pixel Accuracy, Intersection-over-union (IoU) and Dice Coefficient (F1 score) [17] as shown in table 1 with their definitions and calculation formulas. First is pixel accuracy to calculate the proportion of pixels classified correctly of output image which is a straightforward evaluation which higher value means higher segmentation accuracy. However, it does not necessarily imply superior model capability even calculated high accuracy, since the class imbalance issue. Then IoU, also called Jaccard Index, is a well-known and more effective evaluation to assess model. Similarly, higher value means better performance. It is more applicable than pixel accuracy especially for the presence of overlapping regions. Dice Coefficient is positively correlated with IoU which assesses the similarity between prediction and truth. All these three metrics are range from 0 to 1 with larger values means better segmentation performance.

Table 1. Metrics for segmentation

Metrics	Formula
Pixel Accuracy	$\frac{Correct\ number\ of\ pixels}{Total\ number\ of\ pixels}$ Percetage of pixels in correct prediction
Intersection-Over-Union (IoU)	$\frac{Area\ of\ overlap}{Area\ of\ union}$ The ratio of overlap to union between prediction and ground truth
Dice Coefficient (F1 score)	$\frac{2 \times Area\ of\ overlap}{Total\ number\ of\ \ pixels\ in\ both}$ The ratio of double overlap area to total number of pixels of both images

According to the results obtained from the test dataset, the proposed model demonstrates strong performance in surface feature extraction, as indicated by the average pixel accuracy of 0.8542, the average intersection over union (IoU) score of 0.6186, and the average F1 score of 0.7653, the average processing time is 1.4082 seconds. Therefore, the model achieves inteliigent segmentation for different input surfaces in high efficiency. In order to enhance the model's utility for future metrological applications, there is a need to further improve its accuracy, precision, and robustness.

5. Conclusion and future work

In conclusion, this paper applies U-Net model to perform freeform structured surface segmentation task. The experimental results show the remarkable segmentation performance with high accuracy about 85% on feature extraction. Compare to traditional segmentation methods, it can achieve intelligent and automatic segmentation without any manual operations and test phase. Additionally, the use of GPU-supported neural network data processing allows for feature extraction map output within a few hundred milliseconds, indicating high efficiency and low cost. For future investigations, the model can be optimized for better performance by implementing model selection and increasing data diversity. To increase generalisation of model, general engineering surfaces and AM surfaces can be mixed in dataset for training and test to extend applicability of model. The use of conventional edge detection and segmentation methods can serve as a pre-processing step to enhance the labeling quality. Moreover, to more accurately identify features with a particular function, the segmentation output can be combined with a classification task in the subsequent step. Furthermore, it is ideal to combine segmentation network with denoising network to construct an intelligent surface characterisation system for smart metrology development.

References

[1] Gröger, C., Schlaudraff, J., Niedermann, F., & Mitschang, B. (2020). for Advanced Manufacturing Analytics. February 2021, 142–155.

[2] Lazzari, A., Pou, J. M., Dubois, C., & Leblond, L. (2017). Smart metrology: The importance of metrology of decisions in the big data era. IEEE Instrumentation and Measurement Magazine, 20(6), 22–29.

[3] Jiang, X. J., & Whitehouse, D. J. (2012). Technological shifts in surface metrology. CIRP Annals - Manufacturing Technology, 61(2), 815–836.

[4] Jiang, J., & Scott, P. (2020). Advanced Metrology: Freeform Surfaces. (1st ed.) Academic Press Inc.

[5] Lou, S., Pagani, L., Zeng, W., Jiang, X., & Scott, P. J. (2020). Watershed segmentation of topographical features on freeform surfaces and its application to additively manufactured surfaces. Precision Engineering, 63(February), 177–186.

[6] Amoda, N., & Kulkarni, R. K. (2013). watershed segmentation IJETTCS-2013-03-24-035. 2(2), 89–94.

[7] Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image Segmentation Using Deep Learning: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1–22.

[8] Chen, X., Williams, B., Vallabhaneni, S., Czanner, G., Williams, R., & Zheng, Y. (2019). Learning Active Contour Models for Medical Image Segmentation. Computer Vision and Pattern Recognition. https://doi.org/10.1109/cvpr.2019.01190

[9] Lecellier, F., Jehan-Besson, S., & Fadili, J. (2014). Statistical region-based active contours for segmentation: An overview. Irbm, 35(1), 3–10.

[10] Shelhamer, E., Long, J., & Darrell, T. (2017). Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(4), 640–651.

[11] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Lecture Notes in Computer Science, 234–241.

[12] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. HAL (Le Centre Pour La Communication Scientifique Directe).

[13] Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848.

[14] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation,” in 2016 Fourth International Conference on 3D Vision (3DV). IEEE, 2016, pp. 565–571.

[15] Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, PP, 1-1.

[16] Salehi, Seyed Sadegh Mohseni, Deniz Erdogmus, and Ali Gholipour. "Tversky loss function for image segmentation using 3D fully convolutional deep networks." International Workshop on Machine Learning in Medical Imaging. Springer, Cham, 2017.

[17] Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., & Garcia-Rodriguez, J. (2018). A survey on deep learning techniques for image and video semantic segmentation. Applied Soft Computing Journal, 70, 41–65.