

ORIGINAL ARTICLE

Open Access



Three-dimensional reconstruction of industrial parts from a single image

Zhenxing Xu^{1,2}, Aizeng Wang^{1,2*}, Fei Hou³ and Gang Zhao^{1,2}

Abstract

This study proposes an image-based three-dimensional (3D) vector reconstruction of industrial parts that can generate non-uniform rational B-splines (NURBS) surfaces with high fidelity and flexibility. The contributions of this study include three parts: first, a dataset of two-dimensional images is constructed for typical industrial parts, including hexagonal head bolts, cylindrical gears, shoulder rings, hexagonal nuts, and cylindrical roller bearings; second, a deep learning algorithm is developed for parameter extraction of 3D industrial parts, which can determine the final 3D parameters and pose information of the reconstructed model using two new nets, CAD-ClassNet and CAD-ReconNet; and finally, a 3D vector shape reconstruction of mechanical parts is presented to generate NURBS from the obtained shape parameters. The final reconstructed models show that the proposed approach is highly accurate, efficient, and practical.

Keywords Three-dimensional reconstruction, Non-uniform rational B-splines, Industrial parts, Deep learning

Introduction

With the development of intelligent manufacturing, mechanical product production has gradually become increasingly automated, flexible, intelligent, and highly integrated. Thus, artificial intelligence, including three-dimensional (3D) reconstruction and sample data acquisition, is inevitably used. For instance, when manipulators are used for automatic loading and unloading, it is necessary to obtain the 3D data of parts (3D reconstruction) from an image to grasp the object. For parts with irregular surfaces and features that are difficult to measure directly, it is necessary to perform a reverse reconstruction to obtain the size parameters. Current 3D reconstruction methods mostly obtain point cloud

data through 3D scanning and thereafter achieve shape reconstruction by post-processing the point cloud. However, it is difficult to achieve real-time performance and vector reconstruction using this approach. This study proposes an image-based 3D vector reconstruction of typical mechanical products that is highly efficient and can achieve non-uniform rational B-splines (NURBS) based reconstruction with high fidelity, simplicity, and inexpensive consumer cameras.

In the Standard for the Exchange of Product Model Data issued by the International Organization for Standardization, NURBS is the only mathematical method used to define the geometric shapes of industrial products. In digital manufacturing [1], all industrial parts have a unified mathematical expression known as NURBS [2]. In the design and manufacturing processes, NURBS is used not only for computer aided design (CAD) but also for data exchange. For industrial parts, the generation of CAD models depends on the corresponding parameters, such as tooth number, modulus, and tooth width of gears. In general, the parameters depend on the classification of the industrial parts, and the only difference is the value of each parameter (Fig. 1). Therefore, when

*Correspondence:

Aizeng Wang
azwang@buaa.edu.cn

¹ School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China

² Key Laboratory of Aeronautics Smart Manufacturing, Beihang University, Beijing 100191, China

³ State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing 100190, China

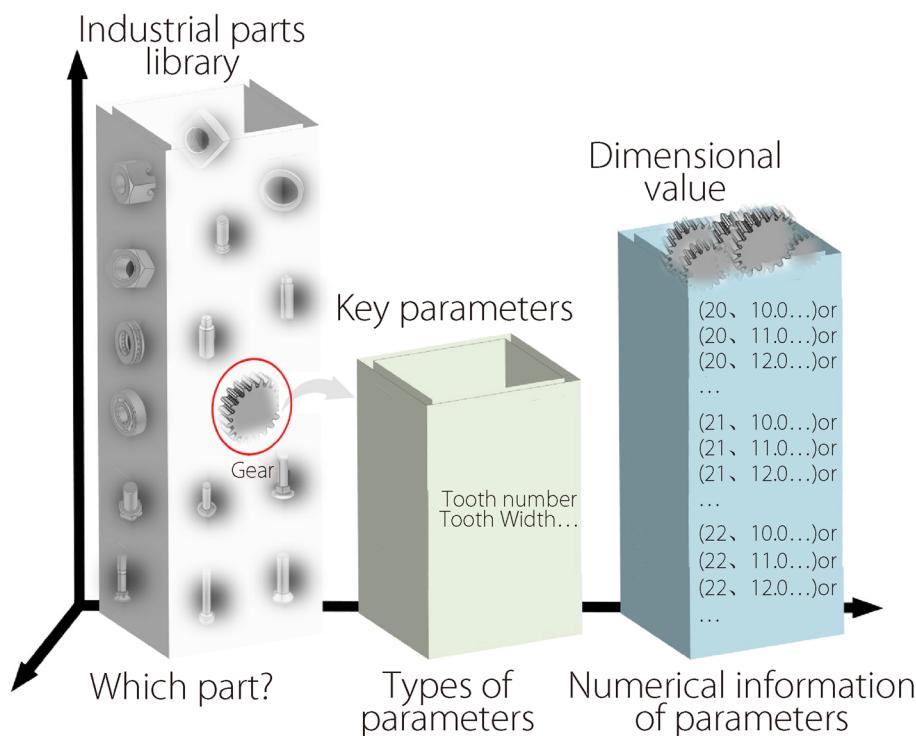


Fig. 1 Parameters of different industrial parts

the type of industrial part is determined, it is possible to reconstruct an accurate 3D model of the part based on NURBS in case its parameters are obtained.

In this study, the research interest was the 3D reconstruction of industrial parts. The contributions of this study are as follows: (1) A dataset of two-dimensional (2D) images of typical industrial parts is constructed, including hexagonal head bolts, cylindrical gears, shoulder rings, hexagon nuts, and cylindrical roller bearings. (2) A deep learning algorithm for parameter extraction of 3D industrial parts that can determine the final 3D parameters and pose information of the reconstructed model using two new nets is developed: a class prediction net (CAD-ClassNet) and a reconstruction prediction net (CAD-ReconNet). CAD-ClassNet was used to determine the type of reconstructed part, and the part parameters were predicted using CAD-ReconNet. (3) A NURBS-based 3D reconstruction of the parts from the parameters obtained by deep learning is presented.

3D reconstruction is classical in computer vision and is widely used in the fields of automatic driving and intelligent robots. Current 3D reconstruction methods based on 2D images can be classified into traditional multiple view geometry approaches [3–5] and deep learning-based methods [6–17]. The former primarily uses a stereo-matching algorithm to recover the 3D structure from a series of 2D images from multiple views obtained by a

camera. However, they cannot recover 3D shapes from a single view. Deep learning-based methods could encode prior knowledge into the network such that they are able to reconstruct the 3D model from a single image. Since AlexNet was first proposed [18], the architecture of deep learning networks has been continuously developing [19–23]. Deep learning has a strong learning ability and good portability, making it easy to achieve excellent results in image classification [20, 21], target detection [19], and image denoising [22]. Deep learning-based single-view 3D stereo methods exhibit better performance than traditional approaches.

Researchers have reconstructed 3D models based mainly on the 2D information fusion of two or multiple views. Jia et al. [7] proposed a dual-view network, DV-NET, which fuses point clouds with two different views using a point-cloud fusion network. Soltani et al. [12] trained data using a depth map and a contour map of multiple views and generated 3D shapes with more details to achieve high-fidelity modeling. Multiple-view-based 3D reconstruction methods have achieved better results; however, it is more challenging to reconstruct 3D shapes from a single image. Single-view-based 3D reconstruction has been applied to buildings [13], furniture [15], human bodies [16], porous media [17], and other structures, particularly indoor furniture. However, they are not applicable to vector model reconstruction,

particularly in intelligent manufacturing and mechanical areas.

The learning ability of a deep learning network relies mainly on a large amount of data. Current 3D reconstruction methods use ShapeNet [24], ObjectNet3D [25], and Pix3D [26] for training. For the aforementioned datasets, the 2D images were aligned with the 3D model using marked points, and different alignment methods were used to improve the reconstruction accuracy. However, it is difficult to fundamentally remove alignment deviations using these methods. Based on MarrNet [11], Sun et al. [26] proposed an approach for shape reconstruction and pose estimation using a 2D–3D alignment dataset. However, these reconstruction methods rely on highly accurate datasets, and it is difficult or expensive to obtain related sample data. Moreover, for shape reconstruction, voxels [27], point clouds [28], and grids [29] have been used to represent reconstructed 3D objects. Although these representations transmit 3D models in a neural network, the final results are not sufficiently accurate without semantic information, and the computation is expensive. Based on NURBS, a unique mathematical representation of industrial products, this study obtained a 3D vector reconstruction of typical industrial parts from a single image. The proposed approach was more efficient, and the final results achieved high accuracy.

Methods

Industrial part dataset generation

For deep learning, its excellent ‘learning’ ability is mostly owing to the training of a large number of samples [30–32]. In practical applications, industry-related historical data can be used to construct training datasets. However, additional industrial sample data are difficult to obtain because of sample acquisition and statistics in additional industries, which hardly meet the large amount of training data required for deep learning. Therefore, obtaining sample sets is closely related to the application of deep learning in industry.

In this study, a 2D image dataset of industrial parts of different sizes and views was constructed, which can be used to construct a feature library of reconstruction parameters. However, it is tedious to obtain a 2D dataset by actual photography, and the accuracy of the camera influences the quality of the sample data, or even the efficiency and accuracy of the final 3D reconstruction. Considering the limitations of actual photography, a CAD model omnidirectional photography approach that can automatically obtain an image dataset of industrial parts with different sizes and poses is proposed in this study.

The size of the input images influences the accuracy of 3D reconstruction and the feasibility of data training in deep learning. The sample sizes are as follows: In this

study, the basic idea of 3D reconstruction was to extract the features of sample images through the employed neural network, and thereafter obtain the 3D parameters through a new algorithm by feature analysis. The relationships between these parameters are as follows:

$$n = s/(0.5 \times a) \quad (1)$$

where s denotes the size of the parts, 0.5 is the proportion of the model in the image, a denotes the accuracy of the parameters, and n denotes the resolution of the image.

The high resolution of sample images will lead to ‘explosion’ of the equipment in data training, and it is difficult for the network to train and fit. Herein, the accuracy of the 3D reconstruction is set to 0.1 mm, and the analysis accuracy of the network is expected to be at the pixel level; thus, the resolution of the input image was 200×200 by Eq. (1) when the model size was 10 mm. The resolution of the input image was 256×256 pixels. Therefore, the size of the original images was set to 256×256 pixels. Additionally, a method to automatically implement the omnidirectional photography of industrial models and successfully obtain a large number of 2D images with a white background is presented. Thereafter, a construction approach for image datasets of industrial parts that is adequate for deep learning in industrial fields is proposed.

To illustrate this further, five typical industrial parts (Table 1) were selected, and each type had ten sizes and 336 poses. Finally, the number of 2D images of the industrial parts dataset was $336 \times 5 \times 10 = 16,800$, where the total shooting points were $7 \times 8 \times 6 = 336$, associated with seven latitude lines, eight longitude lines, and six customary shooting points in the virtual photography space (Fig. 2).

Using the aforementioned parameters and a standard white background, a 2D dataset associated with five typical industrial parts for 3D reconstruction was constructed, as shown in Fig. 3. In this figure, the abscissa axis is sampled from 336 poses of every part, and the vertical axis is sampled from five typical parts, each with 10 sizes. Specifically, “One-Hot” labels are used in the sample image to complete the construction of the dataset. Finally, the dataset is divided into three parts: training, validation, and testing data with ratios of 81%, 9%, and 10%, respectively.

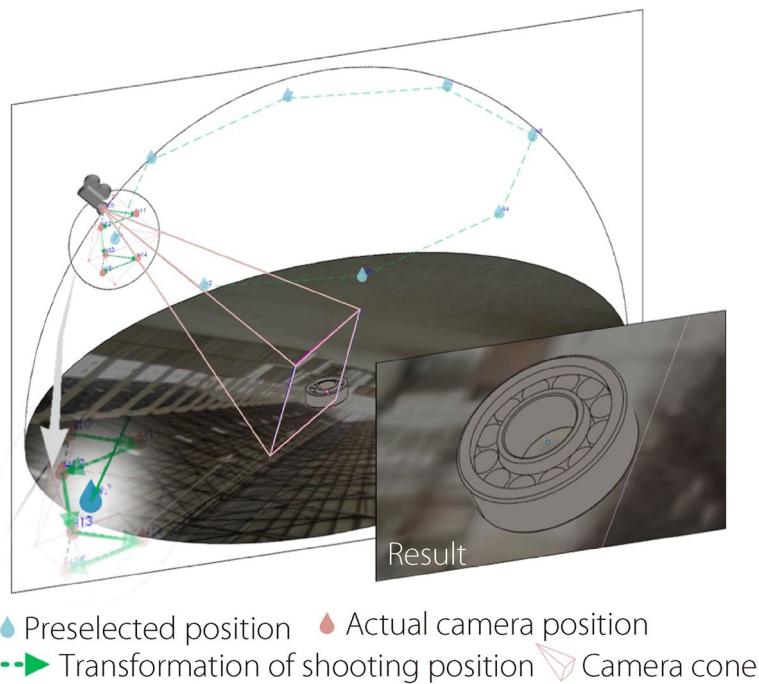
NURBS-based shape reconstruction

This subsection discusses the reconstruction of a 3D shape from a single image and proposes a vector shape reconstruction approach for industrial parts. The proposed method mainly comprises three parts: First, a classification recognition network to distinguish the classes of industrial parts is designed; Second, CAD-ClassNet

Table 1 Five typical industrial parts

| Industrial parts | Parameters | Parameter values |
|----------------------------|--------------------------------|---|
| Hexagon head bolt | [length] | Type _i =[15.00] Type _i =Type ₁ +[1.00]×(i-1) |
| Cylindrical gear | [modulus; tooth width] | Type _i =[1.50, 7.50] Type _i =Type ₁ +[0.10, 0.50]×(i-1) |
| Shoulder ring | [external diameter; thickness] | Type _i =[22.00, 3.25] Type _i =Type ₁ +[1.00, 0.15]×(i-1) |
| Hexagon nut | [nominal diameter; height] | Type _i =[11.00, 8.40] Type _i =Type ₁ +[1.00, 0.76]×(i-1) |
| Cylindrical roller bearing | [external diameter; thickness] | Type _i =[30.00, 10.00] Type _i =Type ₁ +[1.00, 0.33]×(i-1) |

where $i=1, 2, \dots, 10$

**Fig. 2** Virtual photography of industrial parts

and CAD-ReconNet are proposed for industrial parts; Finally, based on CAD-ReconNet, the feature standard library of poses is derived, and the parameters of poses and sizes are obtained from feature analysis.

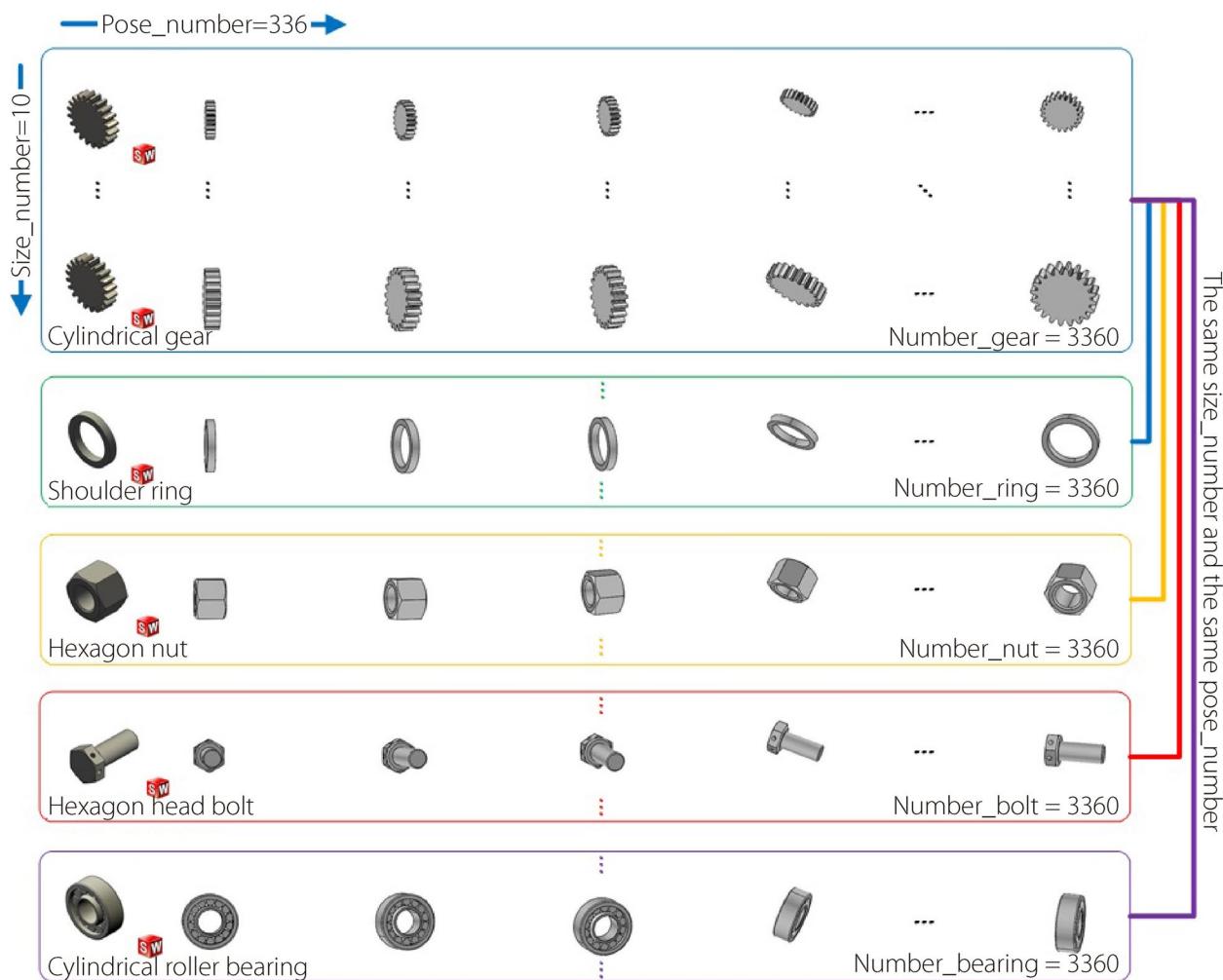
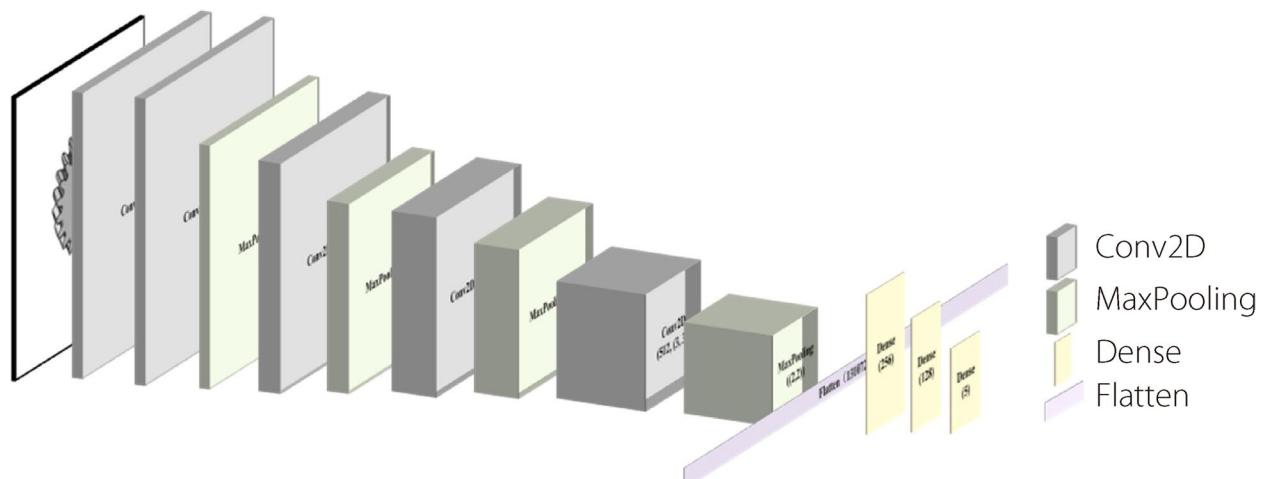
In conclusion, the proposed net comprises two parts: class and reconstruction prediction, where CAD-ClassNet is used to determine the type of the reconstructed part, and the model dimensions can be predicted using CAD-ReconNet.

CAD-ClassNet

To determine the type of reconstructed part, a new network, CAD-ClassNet, was designed, whose structure is

shown in Fig. 4. Five convolution layers and four maximum pooling layers were used to extract image features. The convolution kernel of the five convolution layers was 3×3 and the channel numbers were 64, 64, 128, 256, and 512, respectively. The pool size of the four pooling layers was 2×2 . Thereafter, the 2D output to a one-dimensional vector was flattened and three dense layers were used to complete the classification of the industrial parts.

In CAD-ClassNet and CAD-ReconNet, the softmax activation function in the last layer and the ReLU function in the other convolution layers and full connection layers were used. In particular, the cross-entropy cost function shown in Eq. (2) was used as follows:

**Fig. 3** Standard dataset of five typical parts**Fig. 4** Structure of CAD-ClassNet

$$C = -\frac{1}{n} \sum_x [y \ln a + (1-y) \ln (1-a)] \quad (2)$$

where x denotes the input sample, y denotes the actual label, a denotes the predicted label, and n denotes the total number of samples.

Extraction of poses

Gears are used as an example to represent the concept of shape reconstruction, as shown in Fig. 5. Each class of the five typical industrial parts in this study contains 3360 2D images. The net training and testing processes were completed using this dataset.

The pose determination of parts of 2D images can be regarded as a multiclassification problem. VGGNet and ResNet are often used as feature extractors and classifiers; however, these architectures did not work well on the dataset used in this study. A new network was constructed, CAD-ReconNet (Fig. 5), for pose recognition, in which there are several dense and batch normalization layers. In particular, the net is divided into two parts, a feature extraction block and a feature comparison block, which are described in detail as follows:

Feature extraction block: It comprises seven convolution layers with 3×3 convolution kernels and six maximum pooling layers with a stride of two. Among the seven convolutional layers, the channel number of the first two layers was 64, and that of the subsequent

convolutional layers increased by a ratio of two. In this step, 2048 feature maps with a size of 4×4 were extracted.

Feature compared block: In the second step, the obtained 32,768 dimensional vectors were compressed into 2048, 1024, and 512 dimensional one through one, two, and four full connection layers, respectively. Finally, the input images were classified into 336 poses. In this process, the feature vectors in five different dimensional spaces (32,768, 2048, 1024, 512, and 336) were used as unique features for different poses, and they were also used for pose comparison among different images.

NURBS-based reconstruction

Based on CAD-ReconNet, it is possible to extract the features of the standard dataset of industrial parts and thereafter construct the standard feature library used in pose prediction. The basic idea behind obtaining the dimensions is as follows: (1) Extract the pose feature of the reconstructed 2D image; (2) Compare the pose feature map extracted from the input image with the feature map in the standard feature library and determine the pose information contained in the 2D image; (3) Calculate the similarity between the input image and standard image in the same pose, and thereafter predict the dimensions of the parts according to the similarity. To evaluate the similarity between images, the cosine similarity strategy computed using Eq. (3) was used as follows:

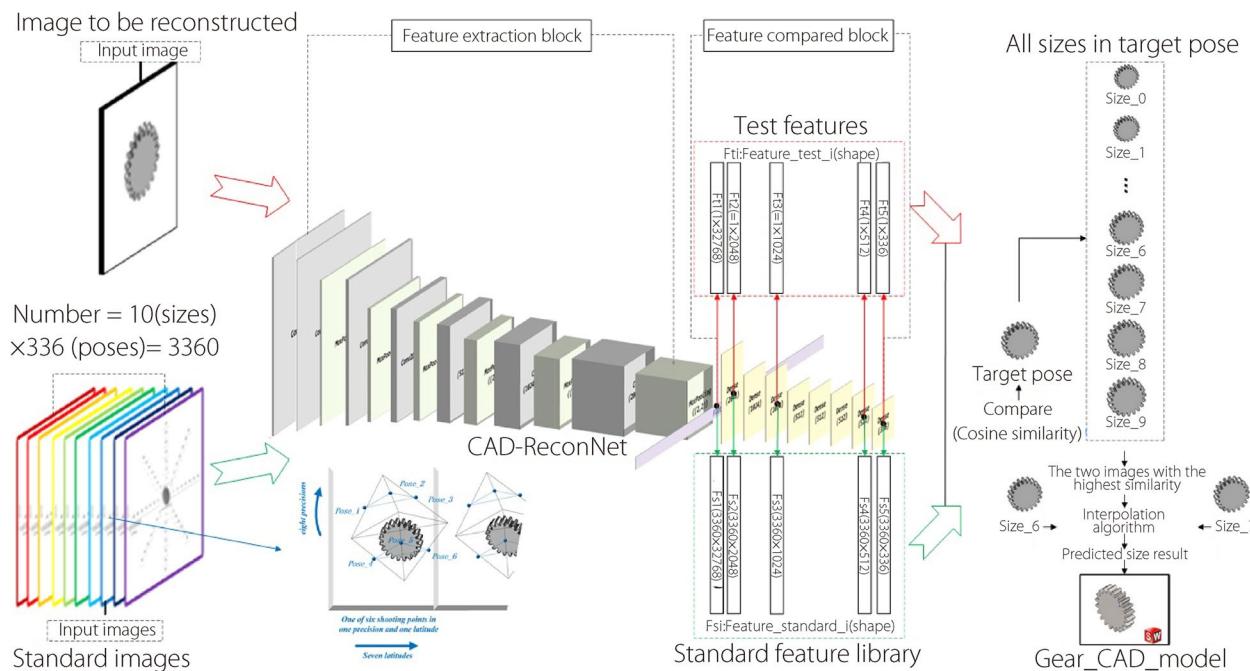


Fig. 5 Extraction of parameters and reconstruction of gears

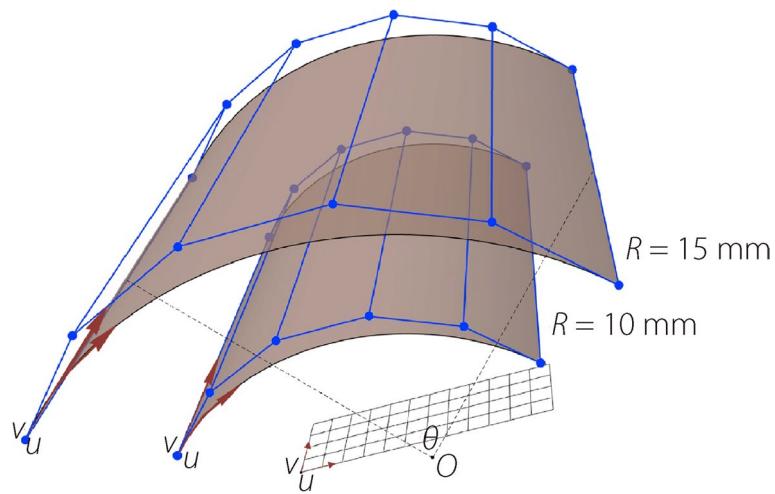


Fig. 6 Scale transformation with same knot vector

$$\cos(\theta) = \frac{\sum_{i=1}^n (A_i \times B_i)}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (3)$$

where A and B denote the image feature vectors. The pose vector dimension is 336.

Two poses from the images associated with the highest similarity were selected as the primary positions.

Thereafter, a subset of standard images with the same position but of different sizes were obtained. By selecting the two sample images associated with the highest similarity, the part size that needed to be 3D reconstructed was predicted. The interpolation coefficients were calculated using Eq. (4), and the final predicted results for the part dimensions were obtained via linear interpolation.

$$Sim_i = \frac{(1-\cos(\theta_{2-i}))}{\sum_i (1-\cos(\theta_i))} \quad (i = 1, 2) \quad (4)$$

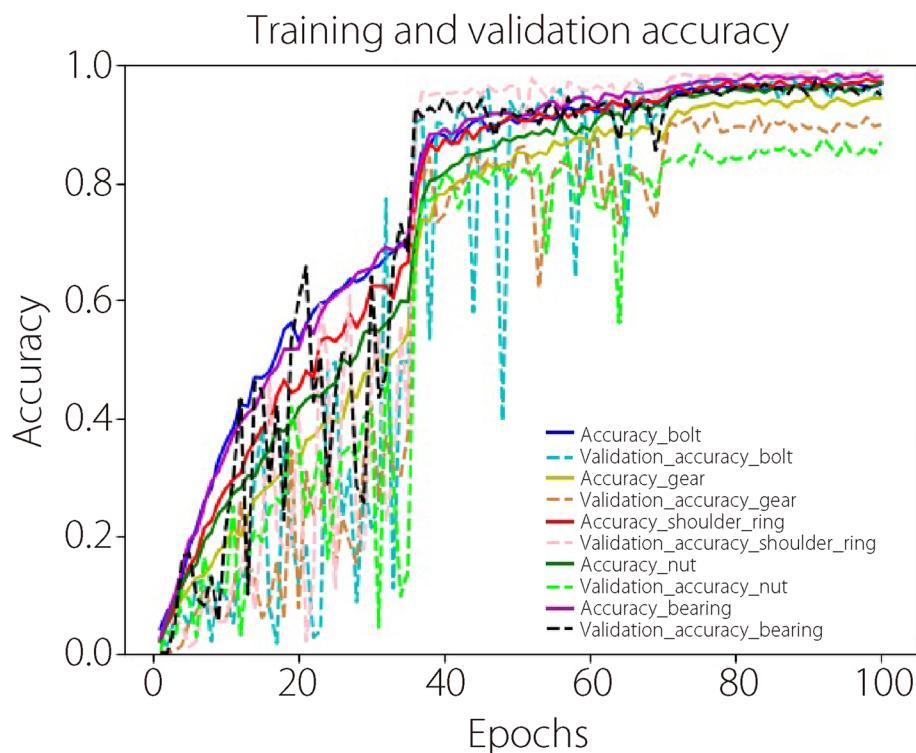


Fig. 7 Accuracy of CAD-ReconNet

Table 2 Performance of CAD-ReconNet training and testing

| Parameters | Minimum training loss | Maximum training accuracy | Minimum verification loss | Maximum verification accuracy | Maximum testing accuracy |
|----------------------------|-----------------------|---------------------------|---------------------------|-------------------------------|--------------------------|
| Hexagon head bolt | 0.1146 | 96.91% | 0.0704 | 97.69% | 97.92% |
| Cylindrical gear | 0.2161 | 94.60% | 0.2239 | 91.75% | 89.29% |
| Shoulder ring | 0.1164 | 97.50% | 0.0444 | 99.34% | 99.70% |
| Hexagon nut | 0.1373 | 96.95% | 0.3928 | 87.46% | 81.85% |
| Cylindrical roller bearing | 0.0782 | 98.49% | 0.0968 | 97.36% | 95.24% |

where $\cos(\theta_i)$ denotes the cosine similarity between images. Furthermore, if the part type is determined, its corresponding control points can be obtained by scale transformation based on the dimension (Fig. 6) and standard model of typical parts. The 3D reconstruction model of the part can be defined using the control points and the same knot vector.

Results and Discussion

Class prediction

To improve the accuracy of class recognition, random rotation, clipping, and bright transformation were applied to the part images for data enhancement. The proposed network has strong convergence and classification ability and achieves good generalization results. Through the testing dataset, the results demonstrate that the classification accuracy of industrial parts is 100%.

Pose prediction

Each part class has its own standard dataset. In the dataset, there are 336 different poses, and each pose has 10

different sizes. In other words, there were several classifications in the dataset, but the number of samples in each classification was relatively small. Batch normalization was used to accelerate the convergence. Strategies for learning the rate gradient decline were used, and the initial value was set to 0.001. A total of 100 epochs were iterated, and the learning rate was changed to 35 and 70 epochs at a ratio of 0.1. The results obtained for this net structure and parameters are shown in Fig. 7.

It is possible to infer that the net loss and accuracy first have large fluctuations because the initial learning rate is set to a larger value (Fig. 7). As the learning rate decreased, the net fluctuation decreased and the net tended to converge. During training and testing, the accuracy and loss of information for each part were monitored (Table 2).

The maximum testing accuracy of the hexagonal nut is slightly higher than 80%, whereas those of the other parts are close to or higher than 90% (Table 2). The strong structural symmetry and small size differences in each direction may explain the low reconstruction accuracy

Table 3 Comparison of CAD-ReconNet and other networks

| Network | Parameters | Maximum training accuracy | Maximum verification accuracy | Maximum testing accuracy |
|--------------|----------------------------|---------------------------|-------------------------------|--------------------------|
| CAD-ReconNet | Hexagon head bolt | 96.91% | 97.69% | 97.92% |
| | Cylindrical gear | 94.60% | 91.75% | 89.29% |
| | Shoulder ring | 97.50% | 99.34% | 99.70% |
| | Hexagon nut | 96.95% | 87.46% | 81.85% |
| | Cylindrical roller bearing | 98.49% | 97.36% | 95.24% |
| ResNet-18 | Hexagon head bolt | 94.56% | 94.72% | 92.56% |
| | Cylindrical gear | 91.25% | 77.89% | 84.52% |
| | Shoulder ring | 85.63% | 84.16% | 85.12% |
| | Hexagon nut | 86.40% | 73.93% | 78.27% |
| | Cylindrical roller bearing | 95.11% | 91.09% | 87.50% |
| ResNet-34 | Hexagon head bolt | 98.31% | 96.70% | 95.54% |
| | Cylindrical gear | 92.65% | 83.83% | 82.44% |
| | Shoulder ring | 98.57% | 92.74% | 92.86% |
| | Hexagon nut | 93.01% | 83.83% | 82.14% |
| | Cylindrical roller bearing | 95.33% | 93.07% | 87.20% |

Table 4 Accuracy of the 3D reconstruction (percentages of every relative error interval)

| Relative error industrial parts | 0%–0.05% | 0.05%–0.1% | 0.1%–0.5% | 0.5%–1.0% | >1.0% |
|---------------------------------|----------|------------|-----------|-----------|-------|
| Hexagon head bolt | 14.73 | 12.50 | 51.04 | 8.78 | 12.95 |
| Cylindrical gear | 42.26 | 25.74 | 18.30 | 1.19 | 12.50 |
| Shoulder ring | 30.06 | 14.29 | 52.64 | 0.15 | 3.87 |
| Hexagon nut | 1.19 | 2.23 | 73.36 | 2.23 | 20.37 |
| Cylindrical roller bearing | 8.78 | 26.34 | 46.58 | 9.38 | 8.93 |

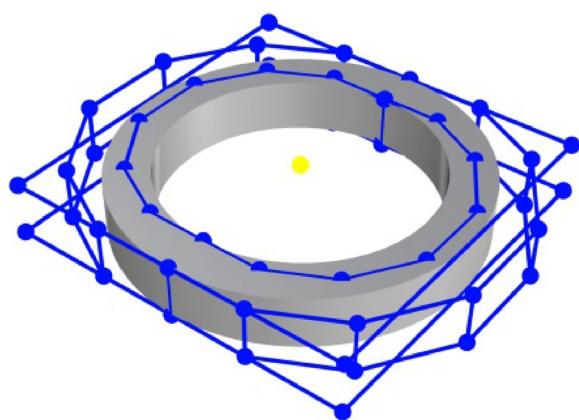
of the hexagonal nut. Overall, it is possible to infer that CAD-ReconNet achieves good results and generalization performance.

CAD-ReconNet was compared with ResNet-18 (batch size=20) and ResNet-34 (batch size=10). Table 3 summarizes the results, and it can be observed that the proposed network converged better. In other words, the features extracted using the proposed network were highly reliable.

Dimension analysis

For industrial parts, the key problem in 3D reconstruction is obtaining the model dimensions, and its accuracy directly affects the validity of the final shape results. Therefore, the accuracy of the size prediction is at the core of net training and testing. For each part, two test sizes (Table 1) were selected and a test set (672 images) was constructed to verify the effectiveness of the proposed method. The proposed approach achieves a better result in size prediction for every typical part (Table 4).

**Fig. 8** Parameter prediction and 3D reconstruction

**Fig. 9** 3D reconstruction of real parts**Fig. 10** NURBS-based geometric modeling of a shoulder ring

From the predicted reconstruction parameters of the typical parts, it is possible to complete the 3D reconstruction based on NURBS from a single image (Fig. 8). To further verify the feasibility of the vector shape reconstruction method, testing images captured from real industrial parts and reconstructed 3D models with high accuracy were used (Fig. 9). Before testing, images from the cameras in the natural scene were preprocessed using background culling and regularization.

According to the results in Figs. 8 and 9, the accuracy of shape reconstruction is less than 0.1 mm from a single image. The proposed system takes an average of 19–20 s to complete the reconstruction of an image with a “NVIDIA GeForce RTX 3070” GPU, an “Intel Core i5-10400F @ 2.90 GHz” CPU, and “16 GB” memory size. Additionally, it outputs the NURBS control points required to define a mechanical part. Thereafter, 3D reconstruction was achieved using geometric modeling based on NURBS. Figure 10 shows the control points and CAD model of the shoulder ring. From these

experiments, the proposed approach can reconstruct industrial parts with high accuracy and efficiency using a single image. Each type of part has known control point structures. Given a model, it is possible to only scale the control edges to fit them to the real dimensions of the part.

Conclusions

In this study, a 3D reconstruction system for industrial parts based on NURBS was constructed, which can achieve the intelligent computation of parameters. Using the predicted parameters, it is possible to reconstruct the corresponding 3D shapes of the industrial parts, which achieves vector reconstruction from a single image. The main contributions of this study are as follows: first, a dataset of 2D images for typical industrial parts is constructed, including hexagon head bolts, cylindrical gears, shoulder rings, hexagon nuts, and cylindrical roller bearings; second, a deep learning algorithm for the parameter extraction of 3D industrial parts is developed using two new nets: CAD-ClassNet and CAD-ReconNet; finally, the 3D shape reconstruction of parts based on NURBS is presented. Examples were provided to illustrate the accuracy and efficiency of the proposed reconstruction approach.

Abbreviations

| | |
|-------|--------------------------------|
| 3D | Three-dimensional |
| 2D | Two-dimensional |
| CAD | Computer aided design |
| NURBS | Non-uniform rational B-splines |

Acknowledgements

Not applicable.

Authors' contributions

AW provided the conceptualization and methodology; ZX, AW, FH and GZ wrote the original draft; AW and ZX reviewed and edited the paper. All authors have read and agreed to the published version of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the Aeronautical Science Foundation of China, No. 2023Z068051002; 2021 Special Scientific Research on Civil Aircraft Project; the Natural Science Foundation of China, Nos. 61572056 and 61872347; and the Special Plan for the Development of Distinguished Young Scientists of ISCAS, No. Y8RC535018.

Availability of data and materials

Not applicable.

Declarations

Competing interests

The authors have no competing interests in the manuscript.

Received: 21 November 2023 Accepted: 5 March 2024

Published: 27 March 2024

References

- Kutin A, Dolgov V, Sedykh M, Ivashin S (2018) Integration of different computer-aided systems in product designing and process planning on digital manufacturing. Procedia CIRP 67:476–481. <https://doi.org/10.1016/j.procir.2017.12.247>
- Piegl L, Tiller W (1997) The NURBS book, 2nd edn. Springer, New York. <https://doi.org/10.1007/978-3-642-59223-2>
- Jin HL, Soatto S, Yezzi AJ (2005) Multi-view stereo reconstruction of dense shape and complex appearance. Int J Comput Vis 63(3):175–189. <https://doi.org/10.1007/s11263-005-6876-7>
- Seitz SM, Curless B, Diebel J, Scharstein D, Szeliski R (2006) A comparison and evaluation of multi-view stereo reconstruction algorithms. In: Proceedings of the IEEE conference on computer vision and pattern recognition, IEEE, New York, 17–22 June 2006. <https://doi.org/10.1109/CVPR.2006.19>
- Hamzah RA, Kadmin AF, Hamid MS, Ghani SFA, Ibrahim H (2018) Improvement of stereo matching algorithm for 3D surface reconstruction. Signal Process: Image Commun 65:165–172. <https://doi.org/10.1016/j.image.2018.04.001>
- Fahim G, Amin K, Zarif S (2021) Single-view 3D reconstruction: a survey of deep learning methods. Comput Graphics 94:164–190. <https://doi.org/10.1016/j.cag.2020.12.004>
- Jia X, Yang SR, Peng YX, Zhang JC, Chen SY (2020) DV-Net: dual-view network for 3D reconstruction by fusing multiple sets of gated control point clouds. Pattern Recognit Lett 131:376–382. <https://doi.org/10.1016/j.patrec.2020.02.001>
- Peng B, Wang W, Dong J, Tan TN (2021) Learning pose-invariant 3D object reconstruction from single-view images. Neurocomputing 423:407–418. <https://doi.org/10.1016/j.neucom.2020.10.089>
- Zhao MH, Xiong G, Zhou MC, Shen Z, Wang FY (2021) 3D-RVP: a method for 3D object reconstruction from a single depth view using voxel and point. Neurocomputing 430:94–103. <https://doi.org/10.1016/j.neucom.2020.10.097>
- Lin CH, Kong C, Lucey S (2018) Learning efficient point cloud generation for dense 3D object reconstruction. In: Proceedings of the 32nd AAAI conference on artificial intelligence, AAAI, New Orleans, 2–7 February 2018. <https://doi.org/10.1609/aaai.v32i1.12278>
- Wu JJ, Wang YF, Xue TF, Sun XY, Freeman WT, Tenenbaum JB (2017) MarNet: 3D shape reconstruction via 2.5D sketches. In: Proceedings of the 31st international conference on neural information processing systems, Curran Associates Inc., Long Beach, 4–9 December 2017
- Soltani AA, Huang HB, Wu JJ, Kulkarni TD, Tenenbaum JB (2017) Synthesizing 3D shapes via modeling multi-view depth maps and silhouettes with deep generative networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, IEEE, Honolulu, 21–26 July 2017. <https://doi.org/10.1109/CVPR.2017.269>
- Yu DW, Ji SP, Liu J, Wei SQ (2021) Automatic 3D building reconstruction from multi-view aerial images with deep learning. ISPRS J Photogramm Remote Sens 171:155–170. <https://doi.org/10.1016/j.isprsjprs.2020.11.011>
- Delanoy J, Aubry M, Isola P, Efros AA, Bousseau A (2018) 3D sketching using multi-view deep volumetric prediction. Proc ACM Comput Graphics Interact Tech 1(1):21. <https://doi.org/10.1145/3203197>
- Yan XC, Yang JM, Yumer E, Guo YJ, Lee H (2016) Perspective transformer nets: learning single-view 3D object reconstruction without 3D supervision. In: Proceedings of the 30th international conference on neural information processing systems, Curran Associates Inc., Barcelona, 5–10 December 2016
- Morales A, Piella G, Sukno FM (2021) Survey on 3D face reconstruction from uncalibrated images. Comput Sci Rev 40:100400. <https://doi.org/10.1016/j.cosrev.2021.100400>
- Feng JX, Teng QZ, Li B, He XH, Chen HG, Li Y (2020) An end-to-end three-dimensional reconstruction framework of porous media from a single two-dimensional image based on deep learning. Comput Methods Appl Mech Eng 368:113043. <https://doi.org/10.1016/j.cma.2020.113043>
- Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th international conference on neural information processing systems, Curran Associates Inc., Lake Tahoe, 3–6 December 2012
- Bochkovskiy A, Wang CY, Liao HYM (2020) YOLOv4: optimal speed and accuracy of object detection. arXiv preprint arXiv: 2004.10934

20. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) [cs.CV]
21. Szegedy C, Liu W, Jia YQ, Sermanet P, Reed S, Anguelov D et al (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, IEEE, Boston, 7–12 June 2015. <https://doi.org/10.1109/CVPR.2015.7298594>
22. Lin K, Li TH, Liu S, Li G (2019) Real photographs denoising with noise domain adaptation and attentive generative adversarial network. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, IEEE, Long Beach, 16–17 June 2019. <https://doi.org/10.1109/CVPRW.2019.00221>
23. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S et al (2014) Generative adversarial nets. In: Proceedings of the 27th international conference on neural information processing systems, MIT Press, Montreal, 8–13 December 2014
24. Chang AX, Funkhouser T, Guibas L, Hanrahan P, Huang QX, Li ZM et al (2015) ShapeNet: an information-rich 3D model repository. arXiv preprint arXiv: 1512.03012
25. Xiang Y, Kim W, Chen W, Ji JW, Choy C, Su H et al (2016) ObjectNet3D: a large scale database for 3D object recognition. In: Leibe B, Matas J, Sebe N, Welling M (eds) Computer vision - ECCV 2016. 14th European conference, Amsterdam, The Netherlands, October 2016. Lecture notes in computer science (Image processing, computer vision, pattern recognition, and graphics), vol 9912. Springer, Cham, pp 160–176. https://doi.org/10.1007/978-3-319-46484-8_10
26. Sun XY, Wu JJ, Zhang XM, Zhang ZT, Zhang CK, Xue TF et al (2018) Pix3D: dataset and methods for single-image 3D shape modeling. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, IEEE, Salt Lake City, 18–23 June 2018. <https://doi.org/10.1109/CVPR.2018.00314>
27. Choy CB, Xu DF, Gwak JY, Chen K, Savarese S (2016) 3D-R2N2: a unified approach for single and multi-view 3D object reconstruction. In: Leibe B, Matas J, Sebe N, Welling M (eds) Computer vision - ECCV 2016. 14th European conference, Amsterdam, The Netherlands, October 2016. Lecture notes in computer science (Image processing, computer vision, pattern recognition, and graphics), vol 9912. Springer, Cham, pp 628–644. https://doi.org/10.1007/978-3-319-46484-8_38
28. Fan HQ, Su H, Guibas L (2017) A point set generation network for 3D object reconstruction from a single image. In: Proceedings of the IEEE conference on computer vision and pattern recognition, IEEE, Honolulu, 21–26 July 2017. <https://doi.org/10.1109/CVPR.2017.264>
29. Wang NY, Zhang YD, Li ZW, Fu YW, Liu W, Jiang YG (2018) Pixel2Mesh: generating 3D mesh models from single RGB images. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds) Computer vision - ECCV 2018. 15th European conference, Munich, Germany, September 2018. Lecture notes in computer science (Image processing, computer vision, pattern recognition, and graphics), vol 11215. Springer, Cham, pp 55–71. https://doi.org/10.1007/978-3-030-01252-6_4
30. Oquab M, Bottou L, Laptev I, Sivic J (2014) Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, IEEE, Columbus, 23–28 June 2014. <https://doi.org/10.1109/CVPR.2014.222>
31. Zhang W, Li X (2022) Data privacy preserving federated transfer learning in machinery fault diagnostics using prior distributions. *Struct Health Monit* 21(4):1329–1344. <https://doi.org/10.1177/14759217211029201>
32. Li X, Yu SP, Lei YG, Li NP, Yang B (2024) Intelligent machinery fault diagnosis with event-based Camera. *IEEE Trans Industr Inform* 20(1):380–389. <https://doi.org/10.1109/TII.2023.3262854>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.