Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Vision Based Topological State Recognition for Deformable Linear Object Untangling Conducted in Unknown Background

Yu Song, Kang Yang and Xin Jiang*
School of Mechanical Engineering and Automation
Harbin Institute of Technology, Shenzhen
HIT Campus Shenzhen University Town, Shenzhen, China
songyu@stu.hit.edu.cn

Yunhui Liu
Department of Mechanical Engineering
The Chinese University of Hong Kong
Hong Kong, China
yhliu@mae.cuhk.edu.hk

***Abstract*— In this paper, we propose a deep learning based method to recognize the topological state of a deformable linear object (DLO). The utilization of deep learning can ensure that topological state recognition is robust to background change. This feature is useful if applications of DLO manipulation in real environment. And this feature has never be realized. In addition, the proposed scheme is also applied to the situation when multiple DLOs exist. This situation has never been considered. By integrating the proposed topological state recognition method and DLO untangling strategy, rope untangling experiments are conducted for both the situations of containing a single DLO and double DLOs.**

I. INTRODUCTION

The development of robot technology is attracting more and more attention from public. At the same time the application fields of robot technology continue to expand. Robots are already helping people to reduce labor cost, increase production efficiency, and improve assembly accuracy. However, the applications of robots in industrial production line are still limited to the operations employing non-deformable rigid objects, such as welding in the production process and assembly of rigid parts. The operation of deformable objects is rarely automatized. In the actual production line, there are still a large number of operations for deformable objects completed manually. For example, in the automobile assembly process, installation of cables still requires manual operation. Besides, cables are still widely used in the power convertor box equipment chassis.

With the increasing demand for robots to operate DLO and the advancement of technology, the research on DLO manipulation is deepened. As for the tracking a DLO, method of fitting using a non-uniform rational B-spline (NURBS) curve [1] and method of fitting simulation model to observation data [2] were proposed. It is also common to use point cloud to obtain the shape of a deformable object and tracking its deformation [3]. In addition, in tracking a deformable object, there is also a way to track the nodes marked on the DLO [4]. There have been some studies on the untangling and knotting problems involved in the operation of DLO [5]. In practical applications, the robotized assembly of cables for automobile production lines is demonstrated in [6]. With the rapid advancement of artificial intelligence, some researchers try to adopt AI in solving the problems in DLO manipulation. A method for learning DLO manipulation through reinforcement learning is proposed in [7] and there is a way of using deep learning to perform retinal vessel segmentation [8].

At present, the topological state recognition of a DLO is mainly obtained by analyzing its projection in a 2D image. In this scheme, status of the crossings including both the overlap and the position of the crossing along the DLO has to be analyzed through image processing. This scheme needs dedicated adjustment to the algorithm when dealing with complicated background. The point cloud processing is often used for this purpose. Recognition of topological state from 3D point cloud represent challenge since the essential information on crossings is not directly demonstrated. With the requirements from extensive applications, it needs a method for recognizing DLO's topological state in more complicated environment.

In this paper, we proposed a method which combines two convolutional neural networks for obtaining the topological state of DLOs. The proposed method is extended to the situations when manipulating multiple DLOs. It is verified that with the proposed deep learning based method, the mutual winding status between DLOs can also be recognized. With the estimated topological status for single DLO and multiple DLOs, untangling strategy is proposed. The advantage of the proposed method is demonstrated by untangling experiments conducted with a robot arm mounted with a dexterous hand. In the experiments, ropes are placed on tables with various complex backgrounds. The successful untangling operations demonstrate that the proposed method is superior than the conventional methods in dealing with the environments with complex backgrounds. This improvement is essential for realizing robotized DLO manipulation in many practical applications.

The remainder of this paper is organized as follows: In Section II, we briefly introduce the recognition process of the topological state and the training process of the neural network. The process of recognizing topological state based on the gradient map is presented in Section III. In Section IV, we describe the proposed untangling strategy and verification experiments in detail. Finally, the conclusions are presented in Section V.

*The corresponding author: Xin Jiang (x.jiang@ieee.org)

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

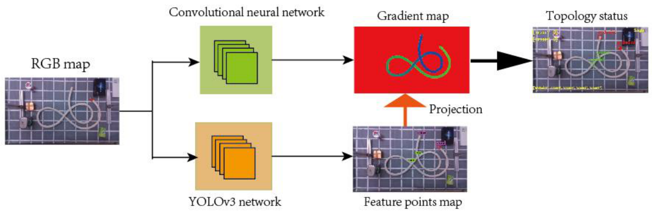
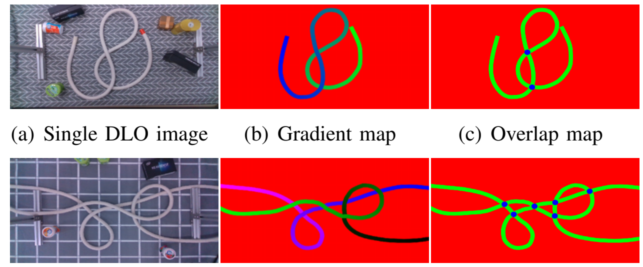


Fig. 1. The processing flow of the proposed recognition method



(a) Single DLO image (b) Gradient map (c) Overlap map

(d) Double DLOs image (e) Gradient map (f) Overlap map

Fig. 2. Dataset for training the network which outputs gradient map

TABLE I

THE PARAMETERS ADOPTED IN NETWORK TRAINING

Project	Parameters
Graphics card model	GTX 1080Ti
Operating system	Ubuntu16.04
Python version	Python 3.6.5
Network framework	Pytorch 0.3.1
Size of the image	640x320
Learning rate	0.0001
Batch Size	2
Number of iterations	600
Number of single training sets	1397
Number of single verification sets	465
Number of double training sets	878
Number of double verification sets	286

II. NETWORK TRAINING

The method for recognizing the topological state of DLO is shown in Fig. 1. It combines two neural networks. One of them is used to detect the crossings in the image containing the DLO and the other one is used to generate the corresponding gradient map. The gradient map here represents the image segmentation result of the scene proposed in [9]. The color variation on the segmented rope segments indicates the segment position along the rope. For each detected crossing point, by referencing its position from the gradient map, we can build the symbol sequence describing the topological status of the DLO.

A. The neural network for generating the gradient map

The acquisition of the gradient map is performed using the deep neural network proposed in [9] and some parameters of the network have been changed to meet the recognition requirements.

For the situation of single DLO, **MSELoss** is used to minimize the error between the original gradient map and the estimated one, which is expressed as:

$$loss(x_i, y_i) = (x_i - y_i)^2. \tag{1}$$

As for the overlap map, **SoftMarginLoss** is chosen, which is expressed as:

$$loss(x, y) = \sum_{n=1}^{N} log(1 + e^{-y_i x_i}) \tag{2}$$

For the situations with double DLOs, **MSELoss** is used in both the component for generating gradient map and the component for generating the overlap map.

B. Dataset preparation and training

In the following experiments, the targets of DLOs are placed on a table with 4 different tablecloths. In addition, on the table there are also bottles and boxes randomly placed. These configurations are to simulate complex backgrounds confronted in real applications. The images for recognition are taken by using RealSense D435 camera. Following the approach in [9], labeled dataset are necessary. Since the approach treats topological state recognition as a segmentation problem solved by deep learning. For each of the image in the dataset, its corresponding label image is the segmentation result in which ropes are painted with gradient color. The color painted on the rope varies gradually when tracing the rope from its one end to the other end regardless of the crossings encountered. The dataset are shown in Fig. 2.

The specification of the training and the main parameters adopted are shown in the TABLE I. Verification conducted by using test data indicates that the network perform well in generating the gradient map. Fig. 3 demonstrates the generated gradient map for a test scene containing one rope. The images in the first row are the original input to the network and the images in the second row are the corresponding gradient maps generated. Fig. 4 represents the training result for scenes containing two ropes. As demonstrated in Fig. 4, in this case, two different ropes are segmented with different color. Particularly, Fig. 5 demonstrates the result obtained by testing the trained network with a scene which is not included in the dataset. It can be seen from the figure that the network can output correct prediction results for the background that has never been seen before. This result verifies that the deep learning based method can help to solve the segmentation problem of DLO from unknown complex background. This special feature is one of our motivations to introduce deep network in topological state recognition.

C. The neural network for cross detection

As for crossing points detection, we choose Yolov3 [10] to realize this function. In order to train the network, 800 images containing single rope and 833 images containing two ropes are prepared as the dataset. The label data is generated by using the tool LabelImg. The training result can be confirmed from Fig. 6.

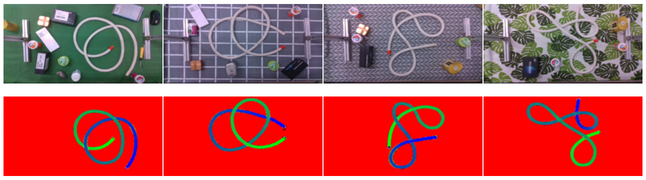


Fig. 3. Predicted gradient map for the scene containing one rope

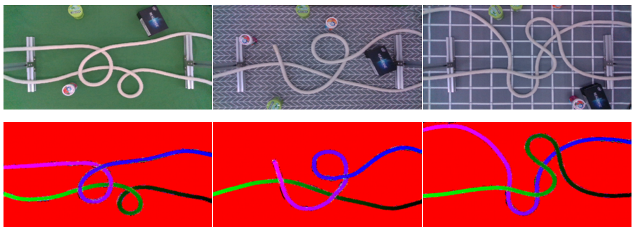
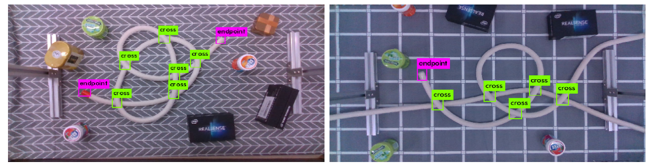


Fig. 4. Predicted gradient map for the scene containing two ropes

(a) (b)

Fig. 5. Gradient map prediction with a test scene



(a) (b)

Fig. 6. Crossing point detection result on test data

III. TOPOLOGICAL STATE RECOGNITION

The aforementioned two neural network constitute the base for topological state recognition. In the following section, we will demonstrate that a simple combination of the results obtained from the two networks can achieve the goal of topological state recognition. First, for every detected crossing, the YOLO network outputs a bounding box which is represented as a region in the image covering the cross. For the information on how the two rope segments overlap, we can refer to the corresponding region in the gradient map. Fig. 7 demonstrates the detected crossings in the gradient map. Thus in the next step it needs to analyze the overlap state by processing the image of gradient map. In the gradient map, boundary lines between segmented areas are generally painted by suspected color and involve more noise than the other areas. The Canny edge detection algorithm [11] is used to extract the boundary line of the intersection area, and then the boundary line and the nearby pixels are removed. Colors in the gradient map except the background color (red) represent the position of the segment from rope's end. In order to better sort the areas by its position (represented by color), the color shown in the gradient map is interpolated between color representing rope' start to rope's end. In this research, the color in the gradient map is represented as the RGB value. Interpolation to the color is realized by weighted sum of the RGB value as shown in (3). The weight value $[\theta_1, \theta_2, \theta_3]$ for single and double DLOs are set separately to [0, 10, 1] and [10, 5, 15]. The weighted sum of W value declines along the rope from its beginning to the end.

$$W = [\theta_1, \theta_2, \theta_3] \cdot [R, G, B]^T \tag{3}$$

In the gradient map, for each segmented area, the average value of W represent its potion along the rope. In addition, the number of pixel within the area indicates whether the segment is above/under the other segment. If the number of the pixel is much more than the other segmented area, then it must be the upper segment. Meanwhile, in the case of double DLOs, we can determine the attribution of each segment by judging the range of W values. The recognition of overlap is conducted in the subsequent process. For each crossing detected on the DLO, two values of W can be determined by analyzing the gradient map. One of the value represents passing this crossing as an upper segment. The value of W is calculated by taking the mean for the above segmented area. The other value of W corresponds to the situation of passing as a lower segment. Its value is calculated by taking the mean for the under segmented areas as shown in Fig. 7. Assume that the gradient map is correctly generated, then the W values for the crossing strictly descend from head to tail of the DLO. Then by sorting the values, the crossing passing sequence when tracing the DLO from head to tail can be obtained. In should be noticed that in the sequence each crossing appear twice. As for the double DLO case, the sorting of crossing is done after determining the attribution of the segments.

With the method described above, the topological state can be recognized. The result for a single DLO is shown in Fig. 8. In the figure, each crossing is listed in the upper left corner. The numbers in parentheses represent the overlap state of the crossing. Assuming to trace the DLO from its beginning to the end, then it will pass through each crossing twice. Here, we call the first passing and the second passing for each crossing. When for the first passing, the crossing is passed as an upper segment, then the number is set as 1. Otherwise, the number is set as 0. The crossings list in the bottom left corner represents the obtained crossing sequence when tracing from beginning to end. As can be seen from the figure, the correct topological state is obtained for the four images. The recognition result for double DLOs is shown in Fig. 9. The two DLOs detected are marked as L1 and L2 respectively. The list in the bottom left corner of the image refers to crossing sequence of the two DLOs. The symbols "X" and "Z" in the image indicate the type of crossing namely mutual intersection or self-intersection. The number after the intersection type symbol is defined with the same meaning in single DLO case.

In practical application scenes, generally the number of

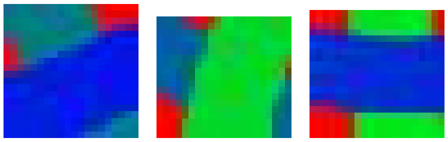


Fig. 7. The cross represented in the gradient map

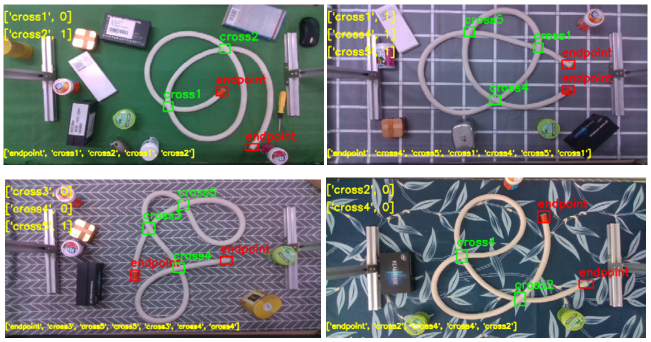


Fig. 8. Topological state recognition result(one single DLO)

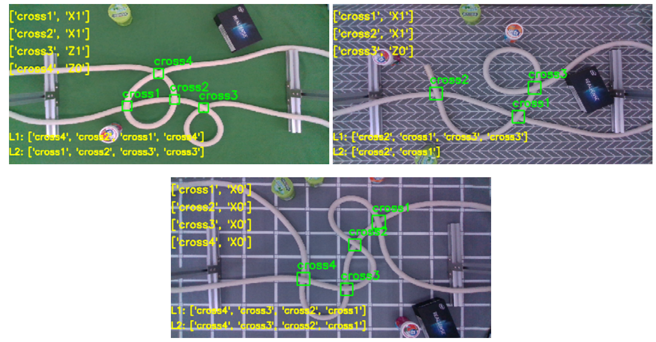


Fig. 9. Topological state recognition result(double DLOs)

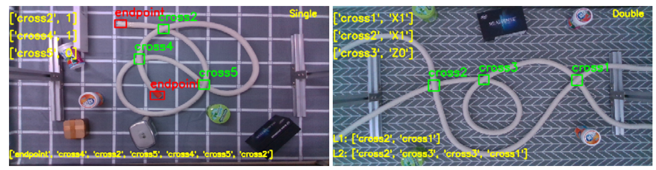


Fig. 10. Recognition for the type of scene

DLOs are not known. It means that the robot has to automatically choose the proper network to recognize the scene since currently the recognition network for single DLO and that for double DLOs are different. We has implement an algorithm to deal with this problem. As the first step of the algorithm, it judges the number of DLOs in the background. For this purpose, we make use of the difference of the scenes containing one DLO and double DLOs. In this research, double DLOs represent the situations where two or more DLOs are tangled together with their end segments outside of the field of view. Thus, for the scenes containing only one single DLO, there will no intersections between the DLO and the boundary of the view. Thus this property is used to distinguish whether there are only one DLO. This is realized by checking whether the segmented DLOs intersect with the boundary of view. As shown in Fig. 10, with the algorithm it can recognize the type of the scene. Then based on the result, the robot can choose to use which model for the subsequent recognition of topological state.

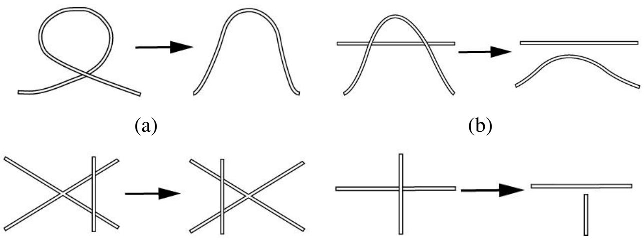
IV. UNTANGLING STRATEGY AND EXPERIMENT

Based on the recognition method, untangling strategy is proposed. It is used in a rope untangling experiments to verify its effectiveness. The details of the experiment are described in this section.

A. *Untangling strategy*

According to the conclusions proved in [12], any topological state for a DLO can be unknotted by sequential combination of four primitive topological transition operation as shown in Fig. 11. Among them, Fig. 11(c) just change the topological state without decreasing crossing number. So we develop an untangling strategy which only utilize the remaining three primitive operations. In order to implement the operation defined in Fig. 11(d), it needs to determine the grasp position on the rope to pull it as shown in Fig. 11. First a ring centered at the crossing is drawn on the gradient map. The radius of the ring is set to 36 pixel and its width set to be 4 pixel. It can be assumed that the ring will intersect with the rope at mainly four places. Two of them are where the end point and the successive segment intersect. The other two places are where the crossing rope segment intersect. The grasp point should be chosen as the point where the successive segment of the end point intersect. After we find all the points where the ring intersect with the rope. The grasp point is determined as the one farthest from the end point. The pulling distance in the program is set as 1.2 times of the distance from grasp point to end-point.

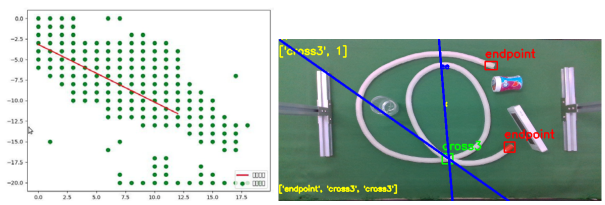
The judgment for the case shown in Fig. 11(b) is done as follows. First it needs to find two adjacent crossings in the sorted crossing list which appear twice. If the overlap state of the two crossings are the same every time they appear, then it satisfies the requirement for applying the primitive operation (b). If we drawn a straight line passing through the center of the two crossings and made it perpendicular to the segment connecting the crossing, then we can obtain two intersecting points. The grasp point for operation (b) is one of the two intersection points. By comparing the color of the points with that of the two crossings, we can determine the positions of the two segments expressed in the sorted crossing list. They are expressed as the first or second segment (means traced for the first time or the second time from the start of the DLO). For the case of single DLO, if the overlap state of the two cross points is estimated as "1" then the “first” segment will be chosen as grasp point, since the segment is the upper one. Otherwise the “second” segment will be chosen as the grasp point. As for the situation with double DLOs, grabbing the first segment if the crossing are all marked as “X1”. When both crossing are marked as “X0”, then the second segment will be chosen. The judgment of the situation shown in Fig. 11(a) is done by checking whether



(a) (b)

(c) (d)

Fig. 11. Four primitive topological transition operations [12]

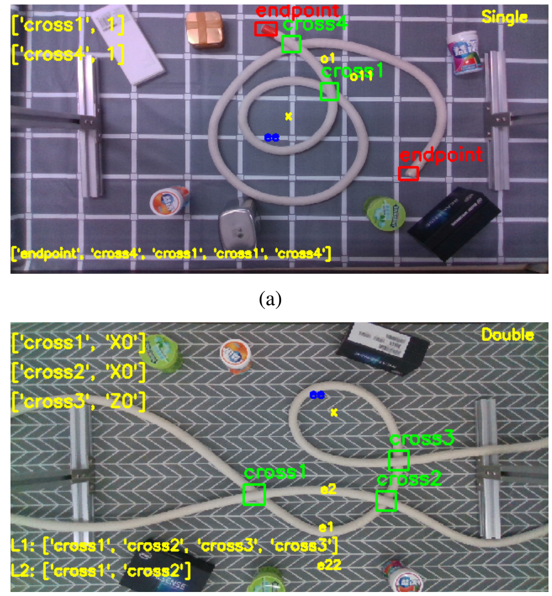


(a) (b)

Fig. 12. The grasp planning in operation

there is a crossing appearing at adjacent positions in the crossing list. The segment cut by the crossing point is fitted as a circle with Hough Circle. If we drawn a straight line by connecting the crossing and the center of circle, then the grasp point is determined as the intersection between the line and the circle. Then for the upper segment of the crossing, another straight line is obtained by fitting. If the angle between this line and the line connecting crossing and center of the circle is less than 90 deg, then the robot and will flip the circle-like segment in anticlockwise direction. Otherwise it segment will be flipped in clockwise direction. The grasp planning in the untangling process is shown in Fig. 12.

Fig. 13 demonstrate the grasp points in untangling process. The yellow "o1" and "o11" in the figure represent the untangling strategies of the situation shown in Fig. 11(d). Among them, "o1" represents the grasp point and "o11" represents the target point. This primitive topological operation can be completed by dragging DLO from "o1" to "o2" to delete the cross point. The yellow "e1" and "e2" in the figure mark the operable points for the first and second segment. The yellow "e22" indicates which one to be operated and the target point. By grabbing the DLO at point "e2", and then dragging to the "e22" point, the primitive topological operation shown in Fig. 11(b) can be completed, and the deleting of the two crossing points in the figure is finished. The yellow "x" in the figure indicates the position of the center of the circle found by the HoughCircles and blue "ee" marks the grasp point for operation. Thus, by grasping the DLO at the "ee" point and rotating according to the calculated rotation direction, the primitive operation of Fig. 11(a) can be completed. The sequence of the primitive operations is set as (b)-(a)-(d).



(a)

(b)

Fig. 13. The operating points in untangling process

(a) (b) (c)

Fig. 14. Dexterous hand grab strategy

B. Primitive operations realized with dexterous hand

In this experiment, we select the dexterous hand of RH8D (Seed Robotics) for rope grasping. The grasping squence with RH8D hand is shown in the Fig. 14. In the experiment, KUKA iiwa7 is used for rope manipulation. For Fig. 11(b) and Fig. 11(d), the robot arm moves to the grasping point first, and then moves to the target point after the dexterous hand finishes the grasping. The robot arm raises the end point position by 100mm during the movement of the grasping point to the target point. It is for avoiding touching other parts of the DLO. For the case of Fig. 11(a), after the target point is grasped, the arm moves up by 100mm and the end point is rotated according to the calculated rotation direction. In the end, the dexterous hand completes the release of the rope.

C. DLO untangling experiment

Finally, the effectiveness of the recognition and untangling strategy is verified by experiments. The experimental platform is shown in Fig. 15. A camera is set to facing down towards the table. It is 800mm away from it.

In the end, we combined the topological state recognition process to the untangling process. Two experiments were conducted. One of them demonstrate the untangling of one single DLO and the other deal with a scene containing two

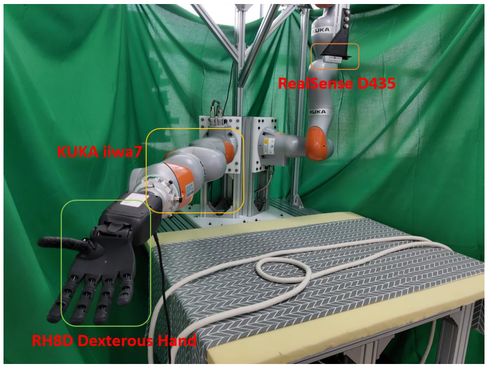
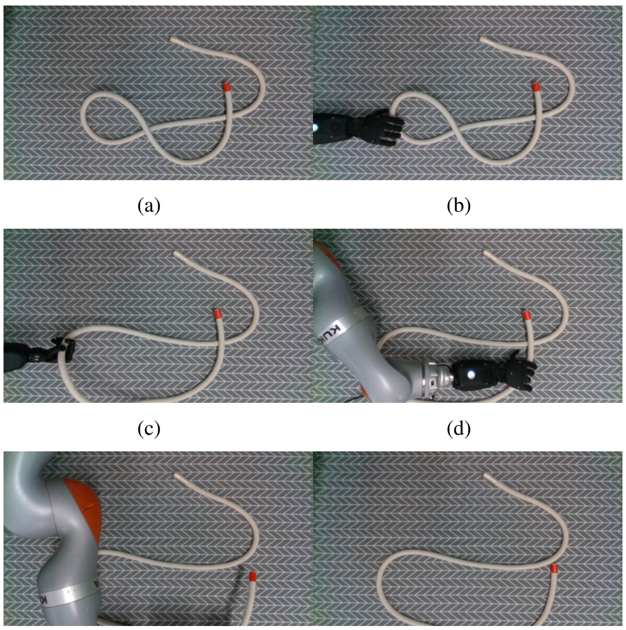


Fig. 15. Experimental platform



(a) (b)

(c) (d)

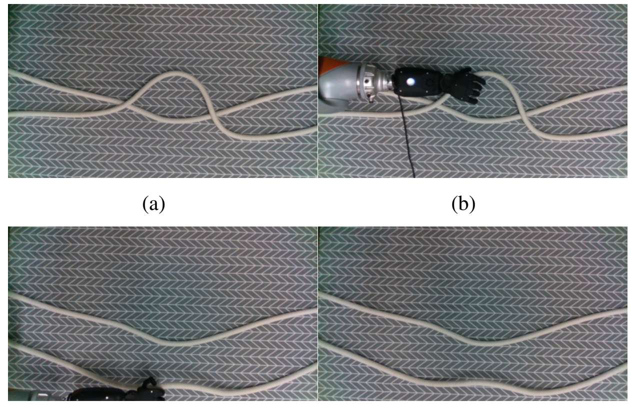
(e) (f)

Fig. 16. Single DLO untangling process

DLOs. It can be seen from Fig. 16 and Fig. 17 that in both experiments the ropes are untangled into simplest state.

V. CONCLUSIONS

In this paper, we proposed a method for recognition and untangling of DLO. First, a gradient map of DLO and the key point position are obtained by convolutional neural network and YOLOv3 network respectively. Then, the results of the two networks are integrated to obtain the topological state of the DLO. We develop a untangling strategy based on the recognized topological state. By implementing the primitive topological transition operation with dexterous hand, untangling of single and double ropes are demonstrated in verification experiments. The two neural networks are innovatively combined and applied to the case of double DLOs. The employment of deep learning in this research can ensure that the method is robust to background change. This feature is useful for applying rope manipulation in environment with complex background.



(a) (b)

(c) (d)

Fig. 17. Double DLOs untangling process

ACKNOWLEDGMENT

This work was supported by the Shenzhen Fundamental Research grant (JCYJ20170811155308088) and National Natural Science Foundation of China (Grant No.61873072), National Natural Science Foundation of China(Grant No. U1713202).

REFERENCES

[1] Jackson R C, Yuan R, Chow D-L, et al. Real-Time Visual Tracking of Dynamic Surgical Suture Threads. in *IEEE Transactions on Automation Science and Engineering*, 2018, 15, pp. 1078-1090.

[2] Javdani S, Tandon S, Tang J, et al. Modeling and Perception of Deformable One-Dimensional Objects. in *IEEE International Conference on Robotics and Automation*, 2011, pp. 1607-1614.

[3] Tang T, Fan Y, Lin H-C, et al. State Estimation for Deformable Objects by Point Registration and Dynamic Simulation. in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 2427-2433.

[4] Lai Y, Poon J, Paul G, et al. Probabilistic Pose Estimation of Deformable Linear Objects. in *IEEE 14th International Conference on Automation Science and Engineering (CASE)*, 2018, pp. 471-476.

[5] Wakamatsu H, Tsumaya A, Arai E, et al. Manipulation Planning for Unraveling Linear Objects. in *IEEE International Conference on Robotics and Automation (ICRA)*, 2006, pp. 2485-2490.

[6] Jiang X, Koo K-M, Kikuchi K, et al. Robotized Assembly of a Wire Harness in Car Production Line. in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 490-495.

[7] Han H, Paul G, Matsubara T. Model-based Reinforcement Learning Approach for Deformable Linear Object Manipulation. in *IEEE Conference on Automation Science and Engineering (CASE)*, 2017, pp. 750-755.

[8] Fu H, Xu Y, Wong D W K, et al. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. in *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 698-701.

[9] Hu Y, Gu Y, Yang J, et al. Multi-stage Suture Detection for Robot Assisted Anastomosis based on Deep Learning. in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 4826-4833.

[10] Redmon J , Farhadi A . YOLOv3: An Incremental Improvement. 2018.

[11] Canny J. A Computational Approach to Edge Detection[J]. in *IEEE Transactions on pattern analysis and machine intelligence*, 1986(6), pp. 679-698.

[12] Wakamatsu, H. Knotting/Unknotting Manipulation of Deformable Linear Objects. in *The International Journal of Robotics Research*, 2006, 25(4), pp. 371-395.