Hao YU, Changjiang ZHOU, Wei ZHANG, Liqiang WANG, Qing YANG, Bo YUAN, 2022. A three-dimensional measurement method for binocular endoscopes based on deep learning. *Frontiers of Information Technology & Electronic Engineering*, 23(4):653-660. https://doi.org/10.1631/FITEE.2000679

A three-dimensional measurement method for binocular endoscopes based on deep learning

Key words: Binocular endoscope; Three-dimensional measurement; Deep learning; Disparity estimation

Corresponding author: Bo YUAN
E-mail: yuanbo@zju.edu.cn
ORCID: https://orcid.org/0000-0002-3185-2690

Motivation

- The conventional endoscope can provide only two-dimensional (2D) images, which are tough to perceive spatial positioning and will eventually affect the accuracy and safety of diagnosis and treatment.

- The binocular endoscope has the ability of three-dimensional (3D) imaging and measurement, which can provide depth information to assist endoscopists in operating the endoscope more accurately, efficiently, and safely [1-4].

- In clinical gastrointestinal endoscopy, the precise size of lesions such as polyps is of great significance for diagnosis. However, the endoscopist's subjective assessment often deviates greatly from the accurate value when operating 2D endoscopes. Therefore, it is necessary to develop an objective, accurate, and efficient 3D measurement method to improve endoscopic surgery results [5].

Main idea

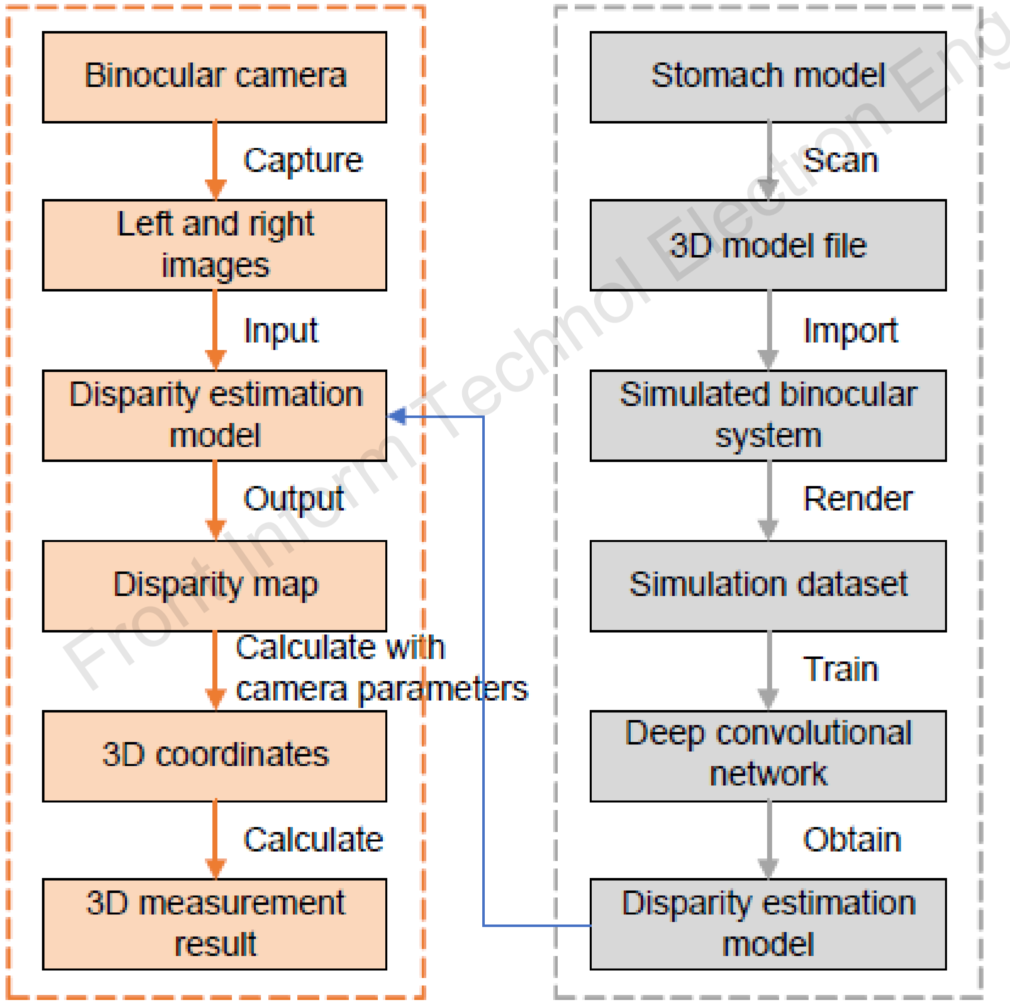
- The binocular matching algorithm based on deep learning can overcome the shortcomings of traditional algorithms, such as semi-global matching (SGM), and will effectively improves the accuracy of disparity predictions in texture-less and high-gloss areas.

- Training a deep learning network requires a lot of data with labels. Given that the ground truth disparity of real binocular endoscopic images is difficult to obtain, we propose a new method for creating simulated binocular images that are very similar to real endoscopic images in terms of content and structure.

- We train an end-to-end disparity estimation network on virtual binocular images and use it to estimate the disparities of real binocular endoscopic images. The 3D measurement results calculated from the estimated disparities are more stable and reliable.

Method

- The overall process of our 3D measurement method based on deep learning is shown in the figure below.



Method

➢ **Create simulated binocular datasets.**

- First, we use a 3D scanner to scan the actual stomach model for obtaining its 3D model file, which is then imported into Blender (an open-source 3D animation production software).
- Then, we can establish a simulated binocular camera system similar to the actual 3D gastroscopic binocular camera in Blender.

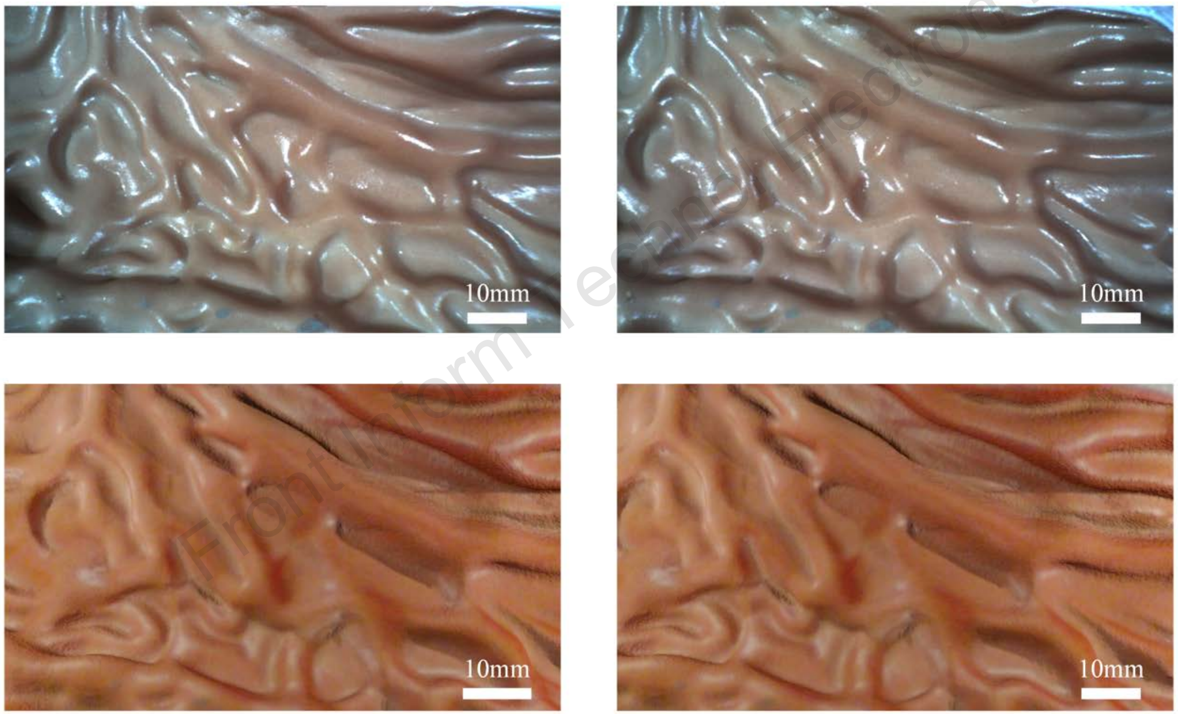
3D scanning

The actual stomach model

The 3D scanned stomach model

Method

- Next, the relative position of the simulated binocular camera to the scanned stomach model is set close to the working state of the actual 3D gastroscope.



The binocular images captured by the actual 3D gastroscope (first row) and binocular images rendered by Blender (second row)

Method

- Finally, the binocular image dataset with ground truth disparity is generated through Blender rendering.

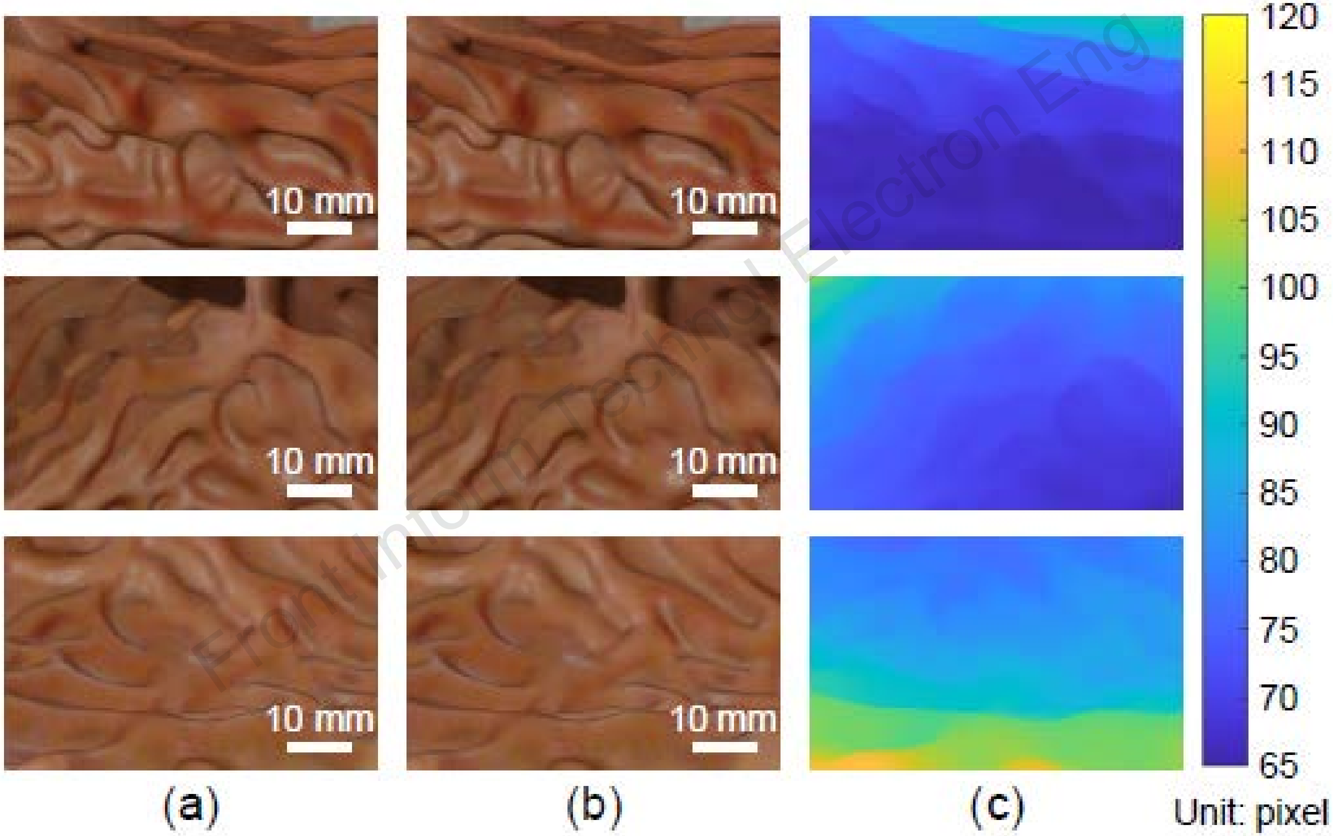
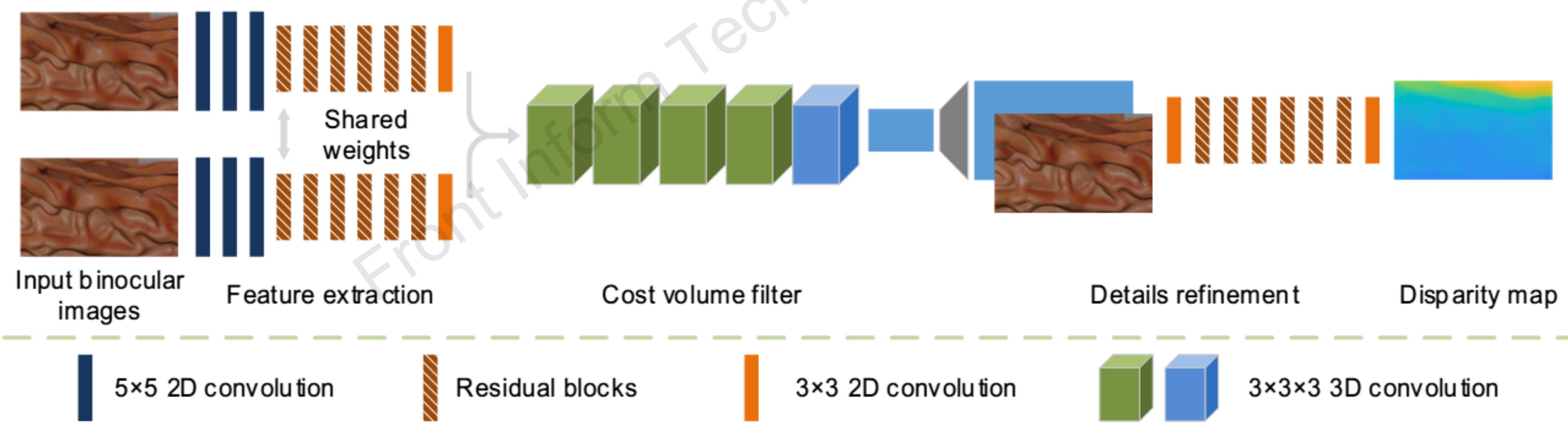


Fig. 3 Examples of the simulation dataset: (a) left camera image; (b) right camera image; (c) ground truth disparity

Method

➢ **Disparity estimation network**

- Our end-to-end disparity estimation network is based mainly on the StereoNet [6] proposed by Google researchers with minor adjustments, which has the advantages of good performance in the weak area, low structure complexity, and high computational efficiency.



Network architecture
The binocular images flow into the network and output the disparity map

Major results

- The disparity map produced by the semi-global block matching (SGBM) algorithm had a lot of noise in inadequate areas such as shadows.
- The disparity map produced by the proposed method was closer to the ground truth disparity map.

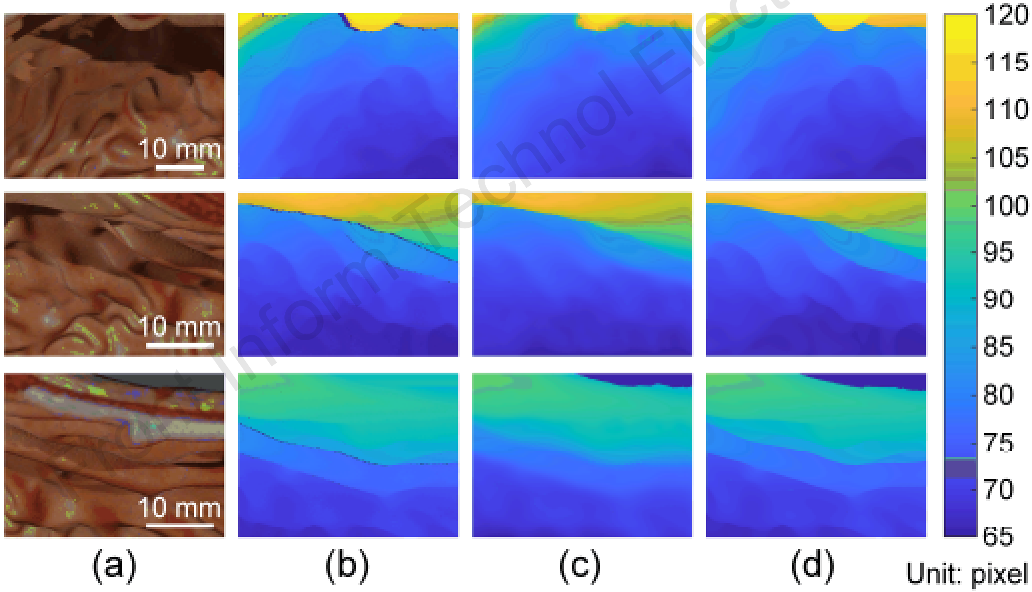


Fig. 6 Disparity results of simulated binocular images: (a) input left image; (b) disparity map produced using the SGBM algorithm; (c) disparity map produced by the proposed method; (d) ground truth disparity map

Major results

- Both the K-pixel-error (KPE) and end-point-error (EPE) [7] of the proposed method have been greatly improved compared to the SGBM algorithm.
- The accuracy measured by EPE increased by 48.9% and the runtime decreased by 90.5%.

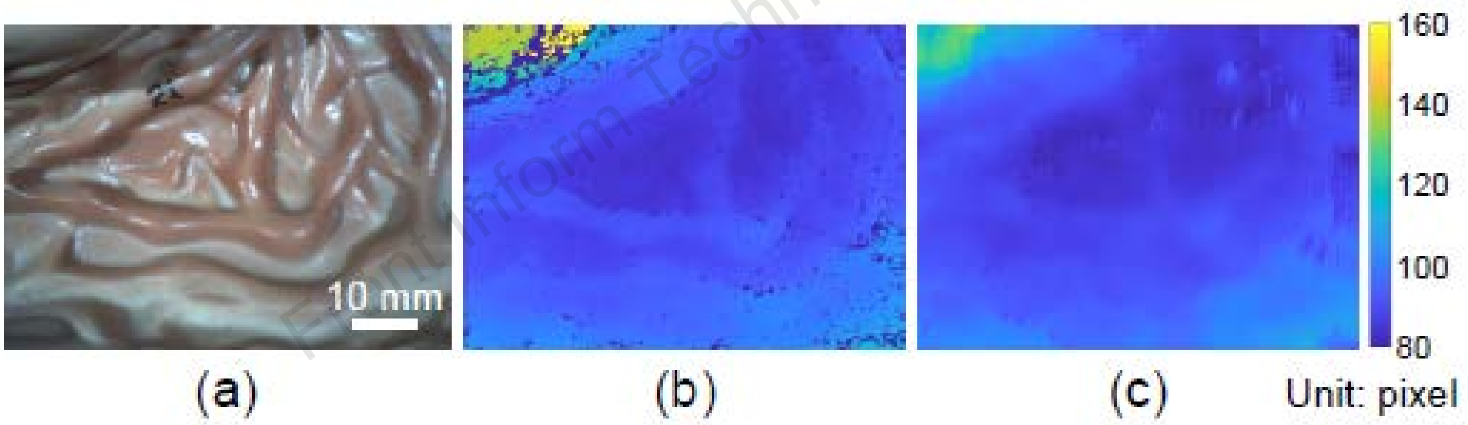
Table 1 Disparity map performance comparison among different methods

Method	KPE (%)			
	5 pixels	7 pixels	9 pixels	50 pixels
SGBM	2.57	2.48	2.42	2.22
Ours	2.20	1.41	1.07	0.46

Method	EPE (pixel)	Runtime (ms)
SGBM	2.25	357
Ours	1.15	34

Major results

- Due to the more complicated imaging conditions in the actual environment, in the disparity map produced by the SGBM algorithm, a large amount of noise was densely distributed in the shadow and overexposed areas.
- The disparity map generated by the proposed method had less noise and was smoother.



Disparity results of real endoscopic images: (a) input image (left); (b) disparity map produced by the SGBM algorithm; (c) disparity map produced by our method

Major results

In addition, the stability of 3D measurement by the proposed method and SGBM was compared. We selected three gastric rugae of the stomach model for measuring the curve length.

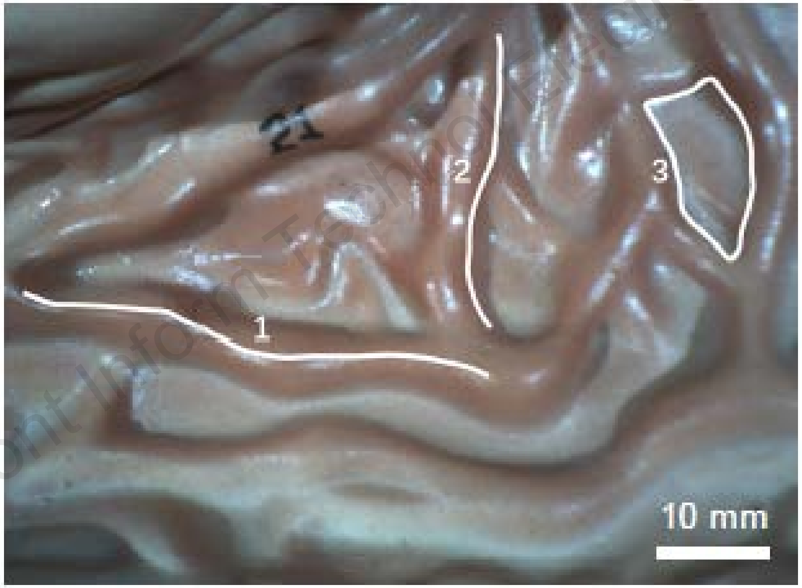


Fig. 8 Three-dimensional measurement curves in the experiment

Major results

The proposed method had no serious errors and the standard deviation was smaller, which demonstrates that the 3D measurement results of the proposed method are more stable and reliable.

Measurement No.	Length (mm)					
	SGBM			Ours		
	Curve 1	Curve 2	Curve 3	Curve 1	Curve 2	Curve 3
1	**14 906.38**	**13 046.76**	44.56	47.09	31.70	53.26
2	46.49	33.96	**14 042.41**	48.24	31.25	53.18
3	44.60	**13 051.37**	47.77	46.85	31.38	51.74
4	**14 920.92**	32.30	**13 693.33**	46.94	31.84	53.07
5	45.60	**13 049.21**	44.28	46.51	31.71	52.46
6	52.70	32.46	**13 774.25**	45.83	31.27	51.99
7	**14 942.16**	32.82	44.07	46.48	31.98	52.88
8	**14 910.01**	**13 046.72**	**14 038.81**	46.77	31.10	53.46
9	43.78	32.76	47.65	47.67	30.76	50.93
10	43.39	34.69	**41 627.01**	46.33	31.65	50.07
Average value (mm)	46.09	33.17	45.67	46.87	31.46	52.30
Standard deviation (mm)	3.43	0.95	1.87	0.69	0.38	1.12

The incorrect measurement results obtained using the SGBM algorithm are in bold and eliminated when calculating the average and standard deviation

Conclusions

- A simulated binocular dataset was generated by establishing a simulated gastroscope scenario through Blender to train a disparity estimation model for 3D measurement.

- The effectiveness of the disparity maps generated from simulated images was evaluated, and the results demonstrated that compared with the SGBM algorithm, the proposed method improved EPE by 48.9% and reduced the runtime by 90.5%. In addition, the proposed method did not create serious errors and had better 3D measurement stability.

- The method developed in this study further promotes the potential application of 3D measurement in the endoscopic environment and can provide accurate and reliable sizes to help endoscopists make accurate diagnoses.

References

[1] Ogino-Nishimura E, Nakagawa T, Sakamoto T, et al., 2015. Efficacy of three-dimensional endoscopy in endonasal surgery. Auris Nasus Larynx, 42(3):203-207.

[2] Cai HF, Wang R, Li Y, et al., 2018. Role of 3D reconstruction in the evaluation of patients with lower segment oesophageal cancer. J Thorac Dis, 10(7):3940-3947.

[3] Nomura K, Kikuchi D, Kaise M, et al., 2019. Comparison of 3D endoscopy and conventional 2D endoscopy in gastric endoscopic submucosal dissection: an ex vivo animal study. Surg Endosc, 33(12):4164-4170.

[4] Omori J, Goto O, Higuchi K, et al., 2020. Three-dimensional flexible endoscopy can facilitate efficient and reliable endoscopic hand suturing: an ex-vivo study. Clin Endosc, 53(3):334-338.

[5] Furukawa R, Aoyama M, Hiura S, et al., 2014. Calibration of a 3D endoscopic system based on active stereo method for shape measurement of biological tissues and specimen. Proc 36th Annual Int Conf of the IEEE Engineering in Medicine and Biology Society, p.4991-4994.

[6] Khamis S, Fanello S, Rhemann C, et al., 2018. StereoNet: guided hierarchical refinement for real-time edge-aware depth prediction. https://arxiv.org/abs/1807.08865v1

[7] Wang XZ, Nie YF, Lu SP, et al., 2020. Deep convolutional network for stereo depth mapping in binocular endoscopy. IEEE Access, 8:73241-73249.