Infrared Image Vein Locating for Robotic IV Insertion

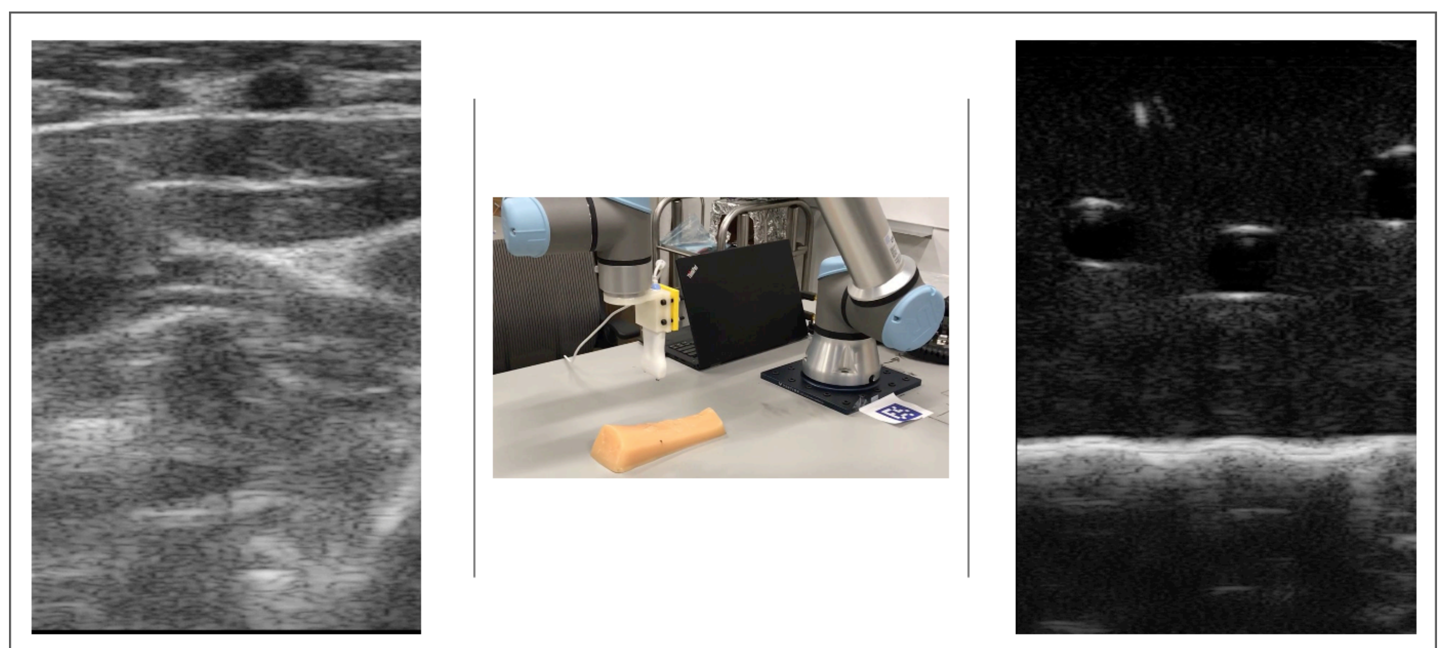
Jing Zhang, Guangsheng Ma, Daniel Buckland
Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA

Abstract

Intravenous (IV) insertion is a medical procedure of inserting a catheter into a vein of patients for delivering medicinal fluids. Robotic IV insertion has been proposed to increase the insertion success rate in recent years. As the frontend module of the robotic platform, vein visual finding plays a critical role in positioning guidance for a human or robot practitioner, and thus is required to behave fast and accurately enough.

Two deep learning models – UNet and YOLOv3 are developed as possible approaches to locate veins on infrared images, which are from two different perspectives – Segmentation and Detection. In this project, the two models are respectively trained on two manually labeled datasets and mainly compared in terms of prediction speed and vein centroid distance, which comes out insights of using segmentation or detection models to deal with the task.

Introduction



Two Datasets:

- Phantom_20: 20 grayscale images of 504 x 747
- Invivo_91: 91 grayscale images of 494 x 754

Methodology

UNet Training Loss Function: pixel-wise binary cross-entropy loss

$$L = \sum_{x\in\Omega} -\left(y\log(p(x)) + (1-y)\,log(1-p(x))\right)$$

YOLOv3 Training Loss Function: summation of basic YOLO loss in three different scales, each of which combines the mean square error on bounding box coordinates, Objectness-Confidence Score and binary cross-entropy loss for class prediction.

$$L = \sum_{i=0}^{S^2}\sum_{j=0}^{B} 1_{ij}^{obj}\left[(x_i-\hat{x}_i)^2 + (y_i-\hat{y}_i)^2 + (w_i-\hat{w}_i)^2 + (h_i-\hat{h}_i)^2\right]$$
$$+\sum_{i=0}^{S^2}\sum_{j=0}^{B}\left[\lambda_{obj}1_{ij}^{obj}\left(-log(p(O_i))\right) + \lambda_{noobj}1_{ij}^{noobj}\left(-log(1-O_i)\right)\right]$$
$$+\sum_{i=0}^{S^2} 1_i^{obj}\, binarycrossentropy(C_i)$$

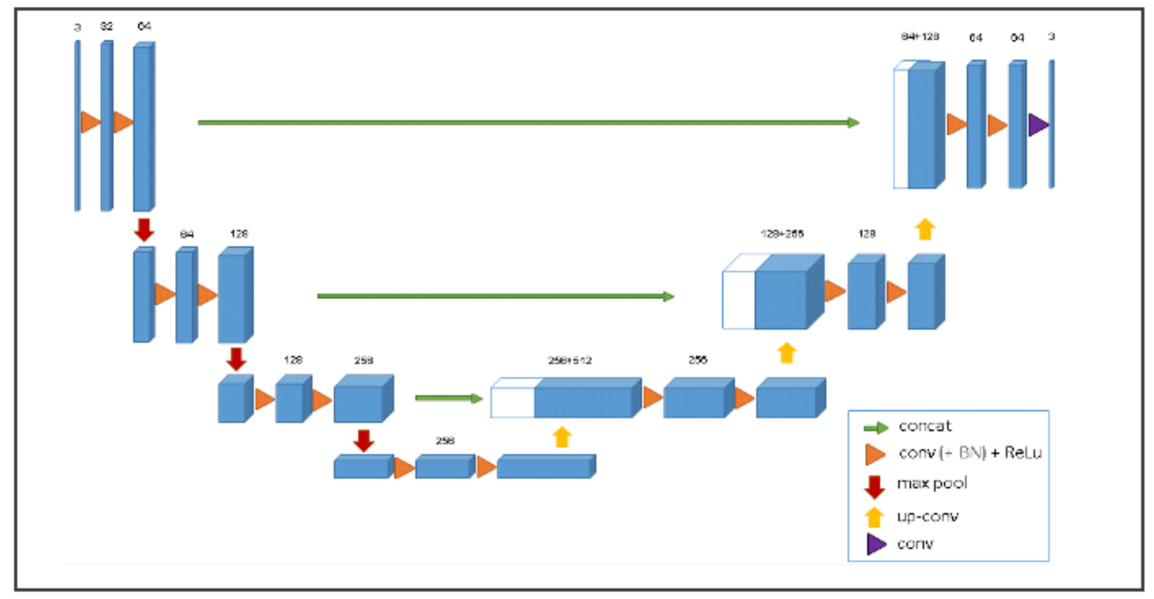


Fig2. Model Structure of UNet

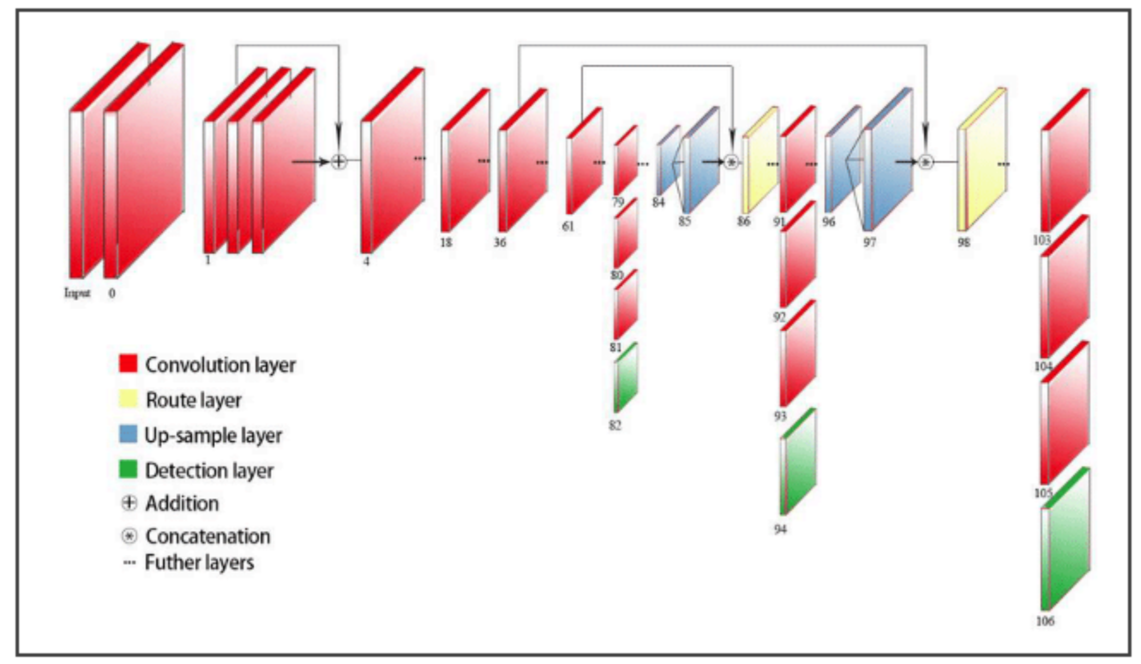


Fig3. Model Structure of YOLOv3
Note: the numbers in the figure shows the layer number in the network.

Results

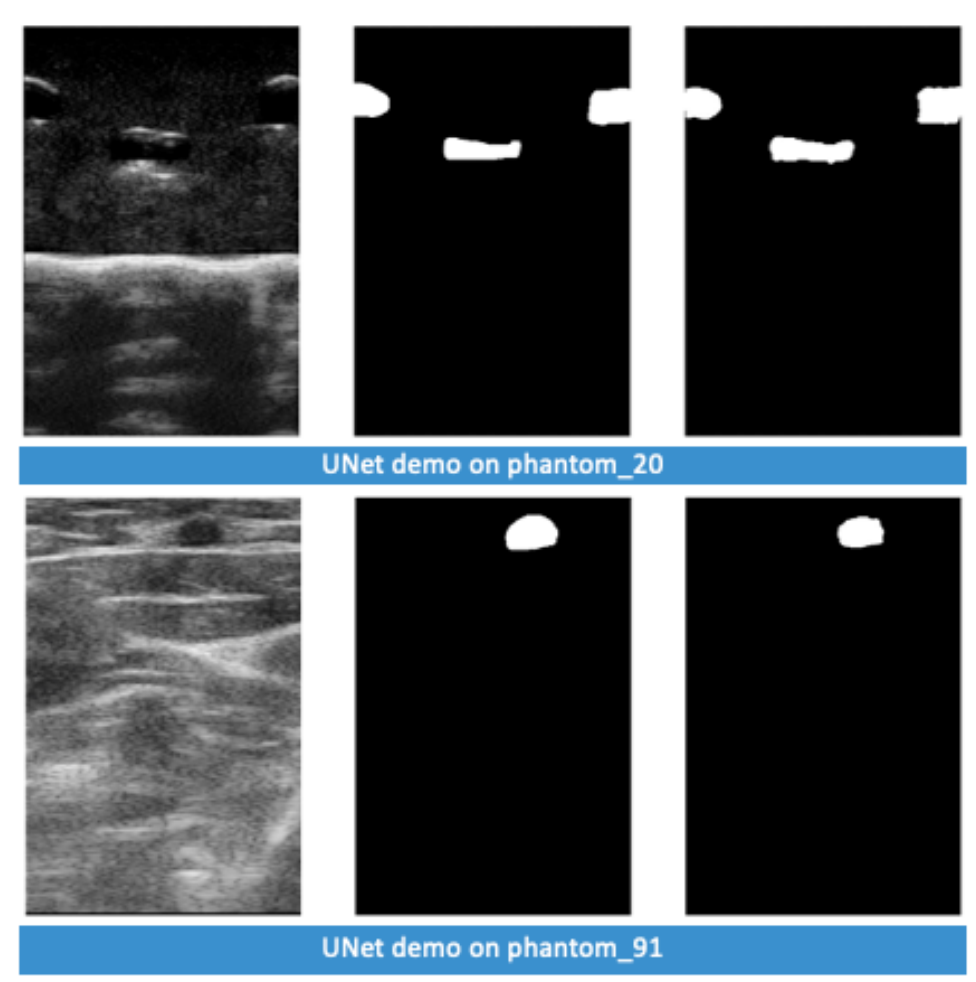


Fig4. Demos of Evaluating UNet on Two Datasets

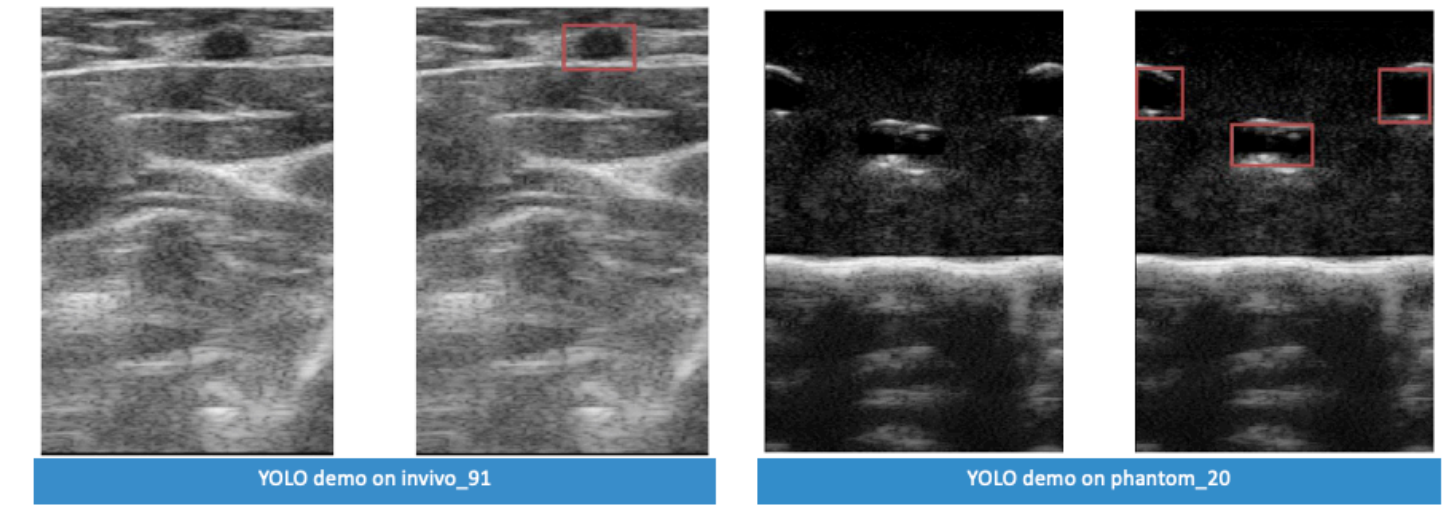


Fig5. Demos of Evaluating YOLOv3 on Two Datasets

Model	UNet		YOLOv3	
Dataset	Phantom_20	Invivo_91	Phantom_20	Invivo_91
Validation Score (IOU)	0.975	0.969	0.780	0.793
Centroid Distance	1.995	4.130	6.736	8.085
Prediction Speed (seconds)	0.081 (gpu) / 2.694 (cpu)	0.023 (gpu) / 2.530 (cpu)	0.025 (gpu) / 0.437 (cpu)	0.024 (gpu) / 0.453 (cpu)

Fig6. Performance Comparison between UNet and YOLOv3
Note: validation score for UNet is dice coefficient; centroid distance means the Euclidean distance between prediction and ground truth; the gpu mentioned above is GeForce RTX 2060.

Analysis:

- UNet and YOLO both locate well where the veins are.
- YOLO's centroid prediction error is obviously greater than UNet.
- Using cpu or gpu does not really affect the prediction accuracy, but affect the prediction speed a lot.

Summary

To conclude, UNet can suggest a more accurate and reliable centroid position of veins on Infrared images while YOLO runs faster than UNet when only cpu is provided and almost same as UNet while using gpu. The robotic IV insertion platform may not choose one solution from them, but rather coordinately put them into real-time operations with further and more comprehensive development.

References

[1] R. D. Brewer and J. K. Salisbury, "Visual Vein-Finding for Robotics IV Insertion." IEEE International Conference on Robotics and Automation, 2010, pp. 4597-4602.

[2] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *ArXiv, abs/1505.04597*.

[3] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

[4] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.