Supplementary Material : Surgical Instrument Segmentation and Self-Supervised Monocular Depth Estimation in Minimally Invasive Surgery: A Multi-task Learning Approach

Stefano Mazzocchetti[1][0009−0001−0104−7586], Laura Cercenelli[1], and Emanuela Marcelli[1]

eDIMES Lab – Laboratory of Bioengineering, Department of Medical and Surgical Sciences, University of Bologna, Via Massarenti, 9, 40138 Bologna, Italy
{stefano.mazzocchett5}@unibo.it

1 Metrics

The metrics employed to assess binary segmentation performance include:

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{1}$$

$$PA = \frac{TP + TN}{TP + FP + TN + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

where TP, FP, TN, and FN denote true positives, false positives, true negatives, and false negatives, respectively.

To assess depth estimation performance, the following metrics are employed:

$$AbsRel = \frac{1}{|D|} \sum_{d \in D} |d^* - d|/d^* \tag{5}$$

$$SqRel = \frac{1}{|D|} \sum_{d \in D} |d^* - d|^2/d^* \tag{6}$$

$$RMSE = \sqrt{\frac{1}{|D|} \sum_{d \in D} |d^* - d|^2} \tag{7}$$

$$RMSElog = \sqrt{\frac{1}{|D|} \sum_{d \in D} |\log d^* - \log d|^2} \tag{8}$$

$$\delta = \frac{1}{|D|} \{d \in D | max(\frac{d^*}{d}, \frac{d}{d^*} < 1.25)\} \times 100\% \tag{9}$$

where d and d^* denote the predicted and ground truth depth values, respectively.

2 Additional Qualitative Results

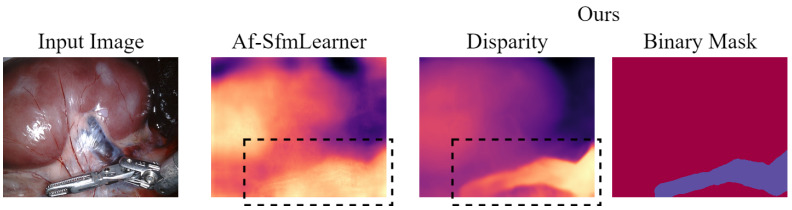

Figure 1 presents zero-shot qualitative results for depth estimation and binary tool segmentation on the Hamlyn Dataset.

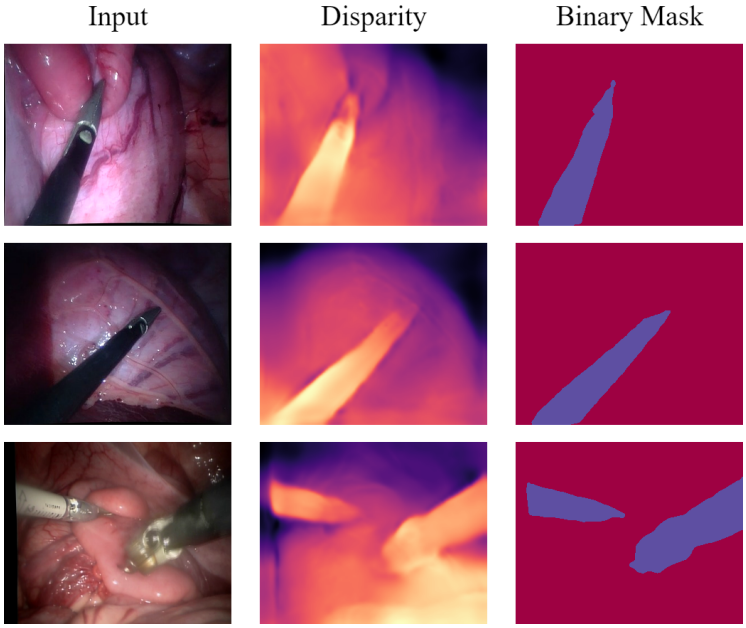

Fig. 1. Zero-shot qualitative results on the Hamlyn dataset.