

Study (Year)	AI Task	Key Performance Metrics
Furube et al. (2025)	Real-time detection of excessive traction on RLN	Correct detection of unintended nerve traction: <b>84.4%</b> ; ETR value correlated with traction degree; AI detected ET earlier than NIM signal in representative case (pre-injury potential)
Furube et al. (2024)	RLN segmentation and recognition metrics	AUC: <b>0.92</b> (left), <b>0.88</b> (right); Dice: <b>0.72</b> ; Sensitivity: <b>0.86</b> ; Specificity: <b>0.89</b> ; IoU: <b>0.40 ± 0.26</b> (right), <b>0.34 ± 0.27</b> (left); surgeon assistance improved RLN recognition rates and IoU with AI assistance
Brandenburg et al. (2023)	Surgomic feature recognition (Active Learning)	Mean F1-score: <b>0.75 ± 0.16</b> (all features); instrument detection F1: <b>0.80 ± 0.17</b> ; inter-rater agreement $\kappa > 0.82$ ; AL improved rare instrument sample selection and performance vs EQS
den Boer et al. (2023)	Anatomical structure segmentation (Bayesian NN)	Median Dice: <b>0.79</b> (azygos/vena cava), <b>0.74</b> (aorta), <b>0.89</b> (lung); algorithm comparable to expert annotations; inference time ~ <b>0.026 s/frame (39 Hz)</b>
Sato et al. (2022)	Recurrent laryngeal nerve (RLN) segmentation	Dice coefficient: 0.58 (AI) vs 0.62 (expert) vs 0.47 (general surgeons); AI performance superior to general surgeons (p = 0.019)
Takeuchi et al. (2022)	Surgical phase recognition	Overall accuracy: <b>84%</b> ; precision: <b>~0.84</b> ; per-phase recall: <b>58–93%</b>