Minimizing Robotic Surgery Adverse Event Through Machine Learning:

TernausNet-11 Implementation with Hamlyn Center 40,000 Robotic Surgery Image Dataset

Gunnar Nelson[1] and Isaiah Herrera[1]

Virginia Tech, Blacksburg, VA, USA[1]

Abstract. Robotic surgery procedures have increased in popularity due to proposed safety and higher patient outcomes. However, the learning curve is a relevant problem in robotic surgery. From past literature reviews, to aid in the learning curve, established training curriculums were created to effectively train surgeons; however, present curriculum gaps have shown surgeons are not trained on how to handle specific adverse events during robotic surgery procedures. With the most frequent problem, having broken pieces fall into patients, further increasing hospitalization costs [2]. The future of machine learning techniques could not only enhance precision and accuracy for surgeons but minimize damage from adverse events. Therefore, to prevent further surgical damage caused from having to recover broken pieces falling into patients, a convolution-neural network (CNN) was fitted with a dataset consisting of 40,000 stereo images of a da Vinci partial nephrectomy, to visually recognize robotic surgery parts, and improve accuracy on a proposed CNN. The CNN is a modified U-NET architecture, the TernausNet-11 created by Dr Vladimir Iglovikov and Alexey Shvets. U-NET architectures are a popular type of CNN, applied for machine learning biomedical imaging. Where, if one were to implement a dataset of certain images, the U-NET architecture performs more convolutions of data, with greater accuracy and fewer data needed. However, the U-NET architecture is constrained for binary imaging. The TernausNet-11 expands on the U-Net, by adding another type of CNN, the VGG-11, which allows coloration of different parts of an image to be implemented [3]. Ultimately, allowing the user the differentiate different parts of an image. However, when the TernausNet-11 was first implemented for the 2017 Robotic Instrument Segmentation Challenge, the accuracy of differentiating different parts of the robotic surgery arm was 45%. The result, according to the study, was due to limited data of 8 videos consisting of 255 frames. Therefore, in order to increase accuracy, a larger dataset was needed. The Hamlyn Centre Laparoscopic / Endoscopic Video Dataset, of a da Vinci partial nephrectomy, contained 40,000 stereo images and when implemented with the TernausNet-11 architecture, increased the accuracy by over 20% [4] [5] [1] [6]. We believe further training of CNNs, in robotic surgery, can not only pave the future for enhancing robotic surgery techniques, but also improve patient outcomes, and lower hospitalization costs.

References

1. Giannarou, S., Visentini-Scarzanella, M., Yang, G.Z.: Probabilistic tracking of affine-invariant anisotropic regions. IEEE transactions on pattern analysis and machine intelligence **35** (03 2012). https://doi.org/10.1109/TPAMI.2012.81
2. Gupta, P., Schomburg, J., Krishna, S., Adejoro, O., Wang, Q., Marsh, B., Nguyen, A., Genere, J.R., Self, P., Lund, E., Konety, B.R.: Development of a Classification Scheme for Examining Adverse Events Associated with Medical Devices, Specifically the DaVinci Surgical System as Reported in the FDA MAUDE Database. Journal of Endourology **31**(1), 27–31 (nov 2017). https://doi.org/10.1089/end.2016.0396, `https://doi.org/10.1089/end.2016.0396`

3. Iglovikov, V., Shvets, A.: Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. ArXiv e-prints (2018)
4. Mountney, P., Stoyanov, D., Yang, G.: Three-dimensional tissue deformation recovery and tracking. IEEE Signal Processing Magazine **27**(4), 14–24 (2010). https://doi.org/10.1109/MSP.2010.936728
5. Pratt, P., Stoyanov, D., Visentini-Scarzanella, M., Yang, G.Z.: Dynamic guidance for robotic surgery using image- constrained biomechanical models. In: Proceedings of the 13th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part I. p. 77–85. MICCAI'10, Springer-Verlag, Berlin, Heidelberg (2010)
6. Ye, M., Johns, E., Handa, A., Zhang, L., Pratt, P., Yang, G.Z.: Self-supervised siamese learning on stereo image pairs for depth estimation in robotic surgery (2017)