

Augmenting Microsurgical Training: Microsurgical Instrument Detection using Convolutional Neural Networks

Tomi Leppanen¹, Hana Vrzakova¹, Roman Bednarik¹, Anssi Kanervisto¹,
Antti-Pekka Elomaa², Antti Huotarinen², Piotr Bartczak², Mikael Fraunberg², and Juha E. Jaaskelainen²

¹School of Computing, University of Eastern Finland, Joensuu, Finland

Email: {tomi.leppanen, hanav, roman.bednarik, anssi.kanervisto}@uef.fi

²Eastern Finland Center of Microsurgery, Kuopio University Hospital, Kuopio, Finland

Email: {antti-pekka.elomaa, ahuotarinen, piotr.bartczak,
mikael.fraunberg, juha.e.jaaskelainen}@kuh.fi

Abstract—In video-based training, clinicians practice and advance their skills on surgeries performed by their colleagues and themselves. Although microsurgeries are recorded daily, training centers are lacking the workforce to manually annotate the segments important for practitioners, such as instrument presence and position. In this work, we propose intelligent instrument detection using Convolutional Neural Network (CNN) to augment microsurgical training. The network was trained on real microsurgical practice videos for which human annotators manually gathered a large corpus of instrument positions. Under challenging conditions of highly magnified and often blurred view, the CNN was capable to correctly detect a needle-holder (a dominant tool in suturing practice) with 78.3% accuracy (F-score = 0.84) with recognition speed above 15 FPS. The result is promising in the emerging domain of augmented medical training where instrument recognition presents benefits to the microsurgical training.

Keywords: Microsurgery, instrument recognition, CNN

I. INTRODUCTION

Watching a brilliant surgeon's video will not make you one, but recognizing what is possible and seeing it done will inspire to be a better surgeon and to achieve the goal.

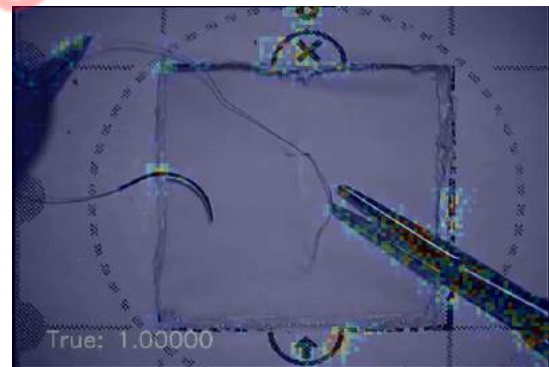
Robert F. Spetzler, June 2016

Microsurgery is considered one of the most demanding fields of medicine. Despite the great needs for effective microsurgical training methods, today's microsurgical training heavily relies on non-interactive video-based materials. While residents are increasingly given access to hands-on practice during real procedures, the availability of such training is limited. Instead, cases recorded in live operating rooms (OR) are typically annotated by expert surgeons and then used as the primary training media. In video-based training, clinicians learn, practice and advance their skills on prior surgeries, performed by themselves or by their colleagues. Skill development in microsurgery, however, cannot be achieved by observations only [1].

Some of the goals of the training are to install bimanual dexterity, increase the speed and accuracy, in order to reduce errors and risks and ensure safety of the patient. Another goal of the microsurgical technical skills training is an objective



(a) Raw view of the operative field



(b) Augmented view with detected instruments

Fig. 1: Real microsurgical procedure during the closing act with a micro forceps (left), a needle-holder (right) and a suture needle (middle). To guide attention of the users, the prototype system automatically detects the needle-holder real time. Heatmap on (b) was created by calculating the gradient of the network.

evaluation of the skill development. Although developed for decades [2] and shown to strengthen expertise [3], the current evaluation frameworks focus merely on the outcome evaluation, as suture quality and execution time.

To augment the microsurgical training, we envision an intelligent system capable of tracking fine movements inside the operation cavity and providing real-time feedback on

the process, training tasks and materials. The functionality has a large potential for future training and intra-operative interactive surgical systems, ranging from automatic annotations of real microsurgeries, studies of surgeon’s eye-hand coordination, safety applications, to building situation awareness models of microsurgeons [4].

Microsurgeries are recorded daily using the infrastructure available at hospitals, however, instrument recognition, detection, and annotation of instrument trajectories are tremendously time demanding. Based on our experiences, complete annotation of a single microinstrument in a 2.5 minutes long video demands at least two hours of focused human attention. Understandably, training centers do not possess workforce, technical skills, and equipment to annotate surgical videos manually.

Existing automatic systems for instrument detection and tracking are unsuitable for microsurgical scenes. There are many reasons why traditional object recognition algorithms fail in this settings. The scenes recorded under the microscope are visually complex, blurred due to unfocused microscope, shaky because of repositioning of microscope, unstable because of tissue handling, and unevenly illuminated, as illustrated in Figures 1 and 2. Furthermore, surgical instruments have non-canonical appearance in general and are only partially visible in the magnified field of view.

To augment human capabilities, advance user experience in watching the surgical and practice videos and to lay groundwork for future mixed- and virtual reality solutions, we propose intelligent instrument detection using Convolutional Neural Network (CNN). Here we investigate the use of deep learning for instrument detection in microsurgery practice videos. Advances in convolutional neural networks (CNN) have been fruitful in various domains as face recognition [5] and object detection [6], with promising overlap to emerging domains as surgical tool tracking.

Our work has the potential to improve learning from surgical and practice materials, increase engagement and direct the focus of the audiences towards the fine orchestra of neurosurgeon’s microinstruments working on the ever-changing but fragile human tissue.

A. Background

The operative field magnified by a surgical microscope poses a challenge for instrument detection and annotation. Visually *restricted field of view*, incorrect *depth perception*, inconsistent *illumination* and pervasive image *blurriness* are inevitable in microscope views. Additionally, the surgical scene is *highly dynamic*. The human tissue on the “background” may be pulsing, bleeding and shaking with micro-tremor caused by the human physiology and repositioning of the microscope, resulting in *motion blur* in the videos.

With no surprise, understanding the operative field requires training of selective attention for a human observer [7], and numerous scene-specific input data for the detection algorithms. Still, traditional object detection algorithms perform poorly when encountering the surgical materials [8].



Fig. 2: Real view on microinstruments. The left column presents well distinguishable instruments, the right column presents the challenging conditions. No post-processing was applied on the frames.

For this reason, a current practice of instrument tracking heavily relies on camera systems and reflexive track points, followed by reconstruction of the trajectories in 3D space. These methods are less of favor among practitioners since the microinstruments gain a significant weight due to the markers and thus, make the training conditions unrealistic [9]. Additionally, motion capture systems often cannot register the fine micro-movements under the microscope.

To detect the instrument trajectories unobtrusively, prior research has attempted to combine various computational techniques. To estimate the position of needle holder in suturing task, Furtado et al. has employed HSV filtering, Watershed algorithm and Harris corner detectors in pipeline for detection, and KLT for optical instrument tracking [8].

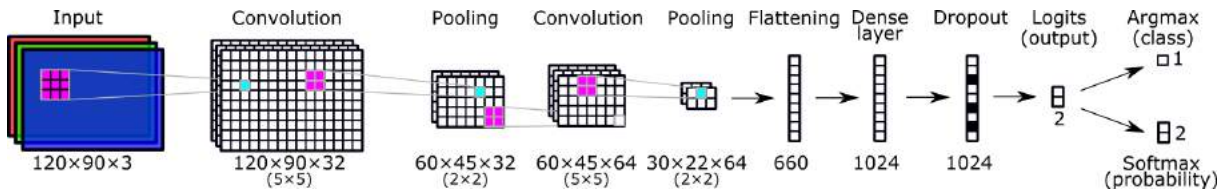


Fig. 3: Architecture of CNN for microinstruments detection.

CNN properties	Value
Parameter initialization	Random
Learning rate	0.001
Network depth	7
Stride size	2×2
Padding	Same
Activation function	ReLU
Pooling method	Max pooling
Optimizer method	Adam optimizer [10]

TABLE I: CNN parameters

Receiving 0.9978 correlation, the proposed methods estimated well a position of a single tool in the simulated and stable laparoscopic environment. Also in the laparoscopy settings, Jiang et al. detected and tracked the tip of instrument using combination of background subtraction and object detection [9]. In comparison to human annotators, their object detection estimated the instrument position with 98.4% accuracy.

In both studies, the simulation environment and available camera systems were dealing with a simplified task under clear image conditions, ideal for object tracking. In this work, we investigate instrument detection in the microsurgical practice scene with realistic instrument movements. Although the scene is partially simplified due to a task board (see Figure 1), instruments movements, in terms of speed, angles and directories, represent the actual microsurgical practice. For automatic recognition, the condition of input frames presents a difficult case for thresholding methods; therefore, deep learning is employed in this work.

Multi-layered neural networks and Deep Learning have gained momentum in numerous object recognition tasks [11]. In the medical domain, Deep Learning and Convolutional Neural Networks, in particular, have been employed in detection, classification, and segmentation of human tissues [12]. In microscopy image analysis, Deep Learning has offered efficient and objective computer-aided diagnosis grounded in raw data [13]. In these cases, however, the input image sets (as in case of brain tumors, breast cancer) present static high-dimensional frames.

Intelligent recognition of instruments and understanding of neurosurgeon's interaction in the magnified field of view is rare. In the endoscopy, Zhang et al. applied CNN for automatic operation scene detection and annotation [14]. Using similarity-based augmentation, CNN could recognize endoscopic instruments (snare tip, forceps tip, and cable) from background frames with 95% accuracy and equal F1 score. The achieved performance has been motivating also

in this study, although the input image characteristics differ fundamentally in the microscopy imaging.

To our best knowledge, we are the first to study instrument detection under high microscope magnification using Convolution Neural Network in microneurosurgical videos.

II. INSTRUMENT DETECTION IN MICROSURGERY

During our joint discussions with the local neurosurgeons, the act of *cutting and suturing* was selected as an compelling candidate for automatic instrument detection. Suturing (e.g. stitching a wound at the end of the surgery) is an important part of neurosurgical toolkit. In *intestinal anastomosis*, for example, a prospective neurosurgeon needs to learn fine hand movements and tool handling under the high magnification of the microscope. Figure 4 illustrates the complicated path of the instrument in one suture.

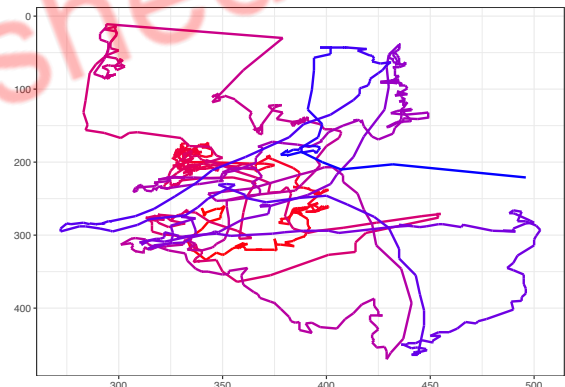


Fig. 4: A trajectory of needle holder during a single suture. Clustered segments present knotting when the needle holder loops around the forceps.

The task of instrument detection splits into following detection subtasks: instrument detection, classification and position localization. In this work, we apply Convolutional Neural Network to investigate instrument detection with additional segmentation for localization. The established CNN is trained on real microsurgical practice videos annotated by human coders.

A. Architecture of Convolutional Neural Network

Convolutional neural network is a multi-layer network, employing convolution to discover patterns in the input image with respect to the labeled output. Our aim is to detect instrument presence in a magnified video. The network

consists of the following layers: *input*, *convolution*, *pooling*, *flattening*, *dense layer*, *dropout* and *output*.

Figure 3 illustrates the structure of our CNN model. In the first layer, CNN accepts frames from the microscope video stream. The input layer is three dimensional (height, width and color channels) and connected to a convolution layer of same size in height and width; however, the third dimension is dedicated to various convolution filters. The value of each tensor of the convolution layer is calculated from the convolution of respective point in the input layer. The connection weights from the previous layer determine the convolution mask. Each of the convolution filters learns a different pattern from the input frame. The convolution layer is connected to the pooling layer which is half the width and the height of the previous convolution layer but retains the filter axis. The value of each tensor is then calculated with maximum pooling; the value is determined by the tensor with strongest connection from its group of tensors in the previous layer.

In our case, we connect two of these convolution layer and pooling layer pairs together so that the second convolutional layer is connected to the first pooling layer. These layers are trained to find patterns in the input layer. The latter pooling layer is flattened to a single dimensional layer and connected to the dense layer. In dense layers, each tensor is connected to every tensor of its previous layer. The weights of these connections are again trained to pick the most useful patterns from the pooling layers to predict the instrument of interest.

In the training phase, we enable dropout in the dense layer. The task of the dropout is to leave some of the tensors out of the training to prevent over-fitting. The dropout is set to keep 40 % of the values of the tensors. The dropout layer is connected to the last - logits- dense layer (length=2). The logits contain likelihoods corresponding to instrument presence in the frame. Using softmax and argmax operations, the likelihoods are transformed into a binary decision.

In training, the order of frames in the dataset is randomized and group individual training samples are split into mini-batches. This settings improves the parallelization and the training speed on a GPU.

B. Dataset: Microsurgical suturing

The original dataset was gathered for practice purposes in collaboration with a local hospital and Neurosurgical Center. In the practice videos, residents were performing sutures in a predefined task board under two magnifications (0.4 and 0.6); the higher magnification increased the task difficulty and allowed the residents to practice for deep microsurgery.

Four microinstruments, as illustrated in Figure 2, were used in the suture practice. A *needle-holder* and a *micro forceps* were hold in the dominant and the non-dominant hand, respectively, interchanging a *needle* in between, when suturing and knotting. When the suture was completed, the needle holder was removed from the view, the micro forceps hold the suture wire and the *microscissors* were employed to cut the loose wire endings. The practice was repeated 12 times, resulting in a recording of 30 minutes approximately.

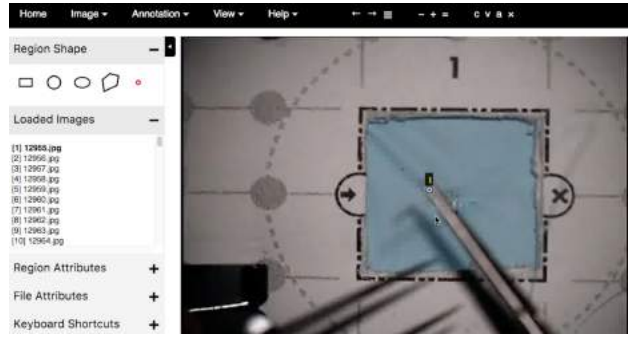


Fig. 5: Annotations of the microinstruments. The human annotators were instructed to annotate the needle holder (right) frame by frame. The instrument is in the dominant hand, carrying the needle during the suturing act, while the micro forceps (left) assists.

13 human annotators annotated the presence and position of *needle-holder* frame-by-frame (as illustrated in Figure 5) in VGG Image Annotator (VIA) [15], with additional post-visualization of quality check. Overall, the annotations of a single tool required 39 hours of human focused work.

When creating a training dataset, we were primarily interested in the needle-holder for its role in the overall resident’s training. With permissions of Center of Neurosurgery, we gathered 57 834 positive and 40 098 negative frames from microsurgical practice videos (97 932). When constructing a training, validation, and test sets, all frames from the videos were gathered and randomly shuffled. The dataset was split into the training (96%), validation (2%) and test set (2%) with unseen frames. The frames in these sets were captured using a built-in Zeiss OPMI microscope camera, recording at 30 Hz with resolution of 720×486 pixels.

Since the training and activation frames might reflect the optics employed in the microscope, we aimed to test how the CNN trained on the field of view from one model of the microscope will perform when the microscope model changes. An additional test set (Test set 2) was recorded using a different microscope (the built-in Zeiss Pico microscope camera), sampling at 30 Hz with resolution of 720×480 . The differences in the resolutions were resolved in the input layer during image downscaling. The suturing practice and instruments employed remained same. Table II summarizes properties of the datasets.

C. CNN Experiments

We use Convolutional Neural Networks (CNN) as basis for image classification and developed two separate ways of utilizing CNNs. First of these approaches, CNN1 in Table III, splits the image into small, overlapping images which are then classified with a CNN. The second approach, CNN2, takes a more traditional path and rescales the whole image into appropriate size (see Figure 3 for the network scheme). Convolutional neural networks are implemented in Python 3.4 using TensorFlow [16].

Both CNNs share the same training video set. For CNN1 we split the frames into images of 128×128 px windows

Video content	Suturing practice under the microscope recorded using Zeiss OPMI and ZeissPico		
Frame rate	30 Hz		
Resolution	712 × 484 (Zeiss OPMI) and 720 × 480 (Zeiss Pico)		
Class labels	Manually annotated by 13 trained annotators		
Training set (Zeiss OPMI)	Whole frame	needle-holder	55 491
		background	38 523
Validation set (Zeiss OPMI)	Frames	needle-holder	1 181
		background	778
Test set 1 (Zeiss OPMI)	Frames	needle-holder	1 162
		background	797
Test set 2 (Zeiss Pico)	Frames	needle-holder	275
		background	226

TABLE II: Properties of training and testing datasets.

with small (2 px) and large (32 px) shifts over the image. This sliding window scans the whole frame area from left to right and from top to bottom with partially overlapping windows. The rationale for the segmentation was to test CNN capabilities in both instrument detection and instrument tip location detection by classifying small parts of the frame at a time. In contrast, CNN2 aims to detect whether the instrument is presented in the frame and is trained and tested on the full frames in the input layer.

The CNN networks were trained using server nodes of a computer cluster Taito at CSC data center with restriction to 24 training hours. The nodes were equipped with four Nvidia Tesla P100 cards and two Xeon E5-2680 v4 CPUs. As TensorFlow can utilize graphics cards for computations, we used one P100 card for TensorFlow model training and six CPU cores for training. The CPU cores were utilized mainly at the beginning of the training process when raw data is extracted from surgical videos. 96 GB of memory was reserved for training on the video data, most of the memory was used to hold downsampled raw video frames for the network. On the other hand, the training on the windowed sample data consumed only 4GB memory.

Human-made annotations were pixel coordinates of the instrument tip. These were converted to binary annotations representing instrument existence. Using this information the annotated frames were randomly labeled to belong to training, validation or test set, so that each set contains approximately the same ratio of positive and negative frames. The binary labels were directly fed to CNN2 that predicts existence of the tool in full frames. In addition, we extracted 128 × 128 pixel windows using the position annotations to create training set for CNN1 that classifies windows. The first extraction saved samples of positive windows that have the tool tip at the center of the window; the second extraction saved negative samples from the background. Finally, malformed samples were removed from the training data.

III. RESULTS

The task in the work was to recognize the presence of the instrument in the microsurgical video. In evaluation of the CNN performance, we measured recognition performance in

terms of accuracy, precision, F-score and processing speed. Since the datasets were approximately balanced, accuracy was employed as the main metric. Table III summarizes the performance of CNN-based detection of the needle-holder on the test (unseen) dataset.

The best performance on the test sets was achieved with the full frames in the activation layer; in experiment E3, CNN2 predicted the presence of the needle holder with 78.3% accuracy (F-score = 0.84) with the recognition speed above 15 fps.

Interestingly, the segmentation of the input frame with 32 px window shift decreased both the performance as accuracy dropped to 67.3% (F-score = 0.78) and the processing speed was almost halved (8.84 fps). The finest window shift (2 px) was tested for instrument localization; CNN was still capable to deliver a performance above the chance level, however, the speed was unsuitable for real-time recognition.

The microscope optics undoubtedly played the important role in the recognition. The evaluation performed on the unseen frames from the second microscope decreased by 12 to 17% in accuracy and 0.07 to 0.22 in F-score, respectively. The results clearly suggest transfer learning will be required to adapt the CNN to particular microscope vendors.

IV. DISCUSSION AND CONCLUSION

Particularly for novice neurosurgeons and residents, everything in the operative field is interesting [17]. Compared to trained experts, novices struggle to direct attention to proper and relevant locations and to focus on the tips of the instruments when needed [7]. In high risk situations, the operative field can get rapidly unreadable due to pulsing brain tissues covered in blood. For practitioners in training, intelligent instrument detection can ease the situation and redirect the focus towards instrument movements. In this work, we presented a first step toward augmented microsurgical video training using CNN.

The best recognition rates (accuracy = 78.3%, F-score=0.84 on the unseen set) were received with CNN trained on full frames without additional segmentation. The recognition performance achieved was in line with the current research. Compared to [14], for example, we have omitted data augmentation to increase sample size in the classes. Interestingly, even without augmentation, CNN trained on whole frames achieved better performance compared to CNN with traditional augmentation. The proposed similarity based augmentation will likely improve the recognition performance and belongs to our future work.

The recognition speed achieved with CNN trained on full frames exceeded 15 fps, suitable for real-time processing. The processing time can be easily improved. The processes for frame extraction and frame saving were implemented with CPUs and covered operations such as resizing and blending; these are suitable candidates for parallelization on GPU cores. Overall, instrument detection is expected to get faster by adding more CPU cores or GPU cards since most components of the process are well parallelizable.

Experiment	Network	Test set	Segmentation	Accuracy	Precision	F-Score	Processing speed [fps]
E1	CNN1	Test set 1	128 × 128, shifted by 32 px	0.673	0.650	0.784	8.864
E2	CNN1	Test set 1	128 × 128, shifted by 2 px	0.601	0.600	0.750	0.066
E3	CNN2	Test set 1	Full frame	0.783	0.735	0.846	15.186
E1	CNN1	Test set 2	128 × 128, shifted by 32 px	0.553	0.551	0.710	8.789
E2	CNN1	Test set 2	128 × 128, shifted by 2 px	0.551	0.550	0.710	0.066
E3	CNN2	Test set 2	Full frame	0.519	0.546	0.625	17.893

TABLE III: Detection rates on the testing sets.

Finer segmentations delivered lower performance. The finer window shift could have introduced more possibilities for false predictions. Additionally, interaction with tissues and instruments under the microscopy has been known to be prone to depth distortion and motion blur [7]. As systematically evaluated by Dodge and Karam, currently available deep learning models are highly impacted by image quality, especially by blur and noise [18]. In future work, we will follow Dodge and Karam’s recommendation to train CNN specifically targeting blurry microscope input.

In future work, instrument detection will benefit from advanced neural network concepts, for example *transfer learning*, *multi-object learning* and *end-to-end learning* [11]. Each of the concepts solves a different obstacle. Training and re-training of a whole network is tremendously time demanding. Transfer learning will potentially use an existing larger pre-trained networks with pre-trained low-level features and adapt the last layers to microsurgical settings.

Multi-object deep learning presents an expected extension of the current CNN as majority of microsurgeries are performed with several tools at once. Finally, multi-instrument detection and localization is too complex for a single classifier; therefore, end-to-end learning will allow us to split the problem to different recognition pipelines and utilize the result in the end.

A. Medical applications and future work

Instrument detection offers numerous emerging applications both in realistic surgical training and in live neurosurgeries in OR. In this work, we envisioned residents and microsurgeons in training who learn not only from pre-recorded video surgeries but also from regular visits at OR where the microscopic view is duplicated at side screens. In these contexts, real-time instrument augmentation allows for better understanding of expert actions and for faster learning of expert’s skills.

In the OR teams, responsible for assisting the neurosurgeons with the right instrument at right time, instrument detection presents a computer-mediated technique for improving interaction between the neurosurgeon and the team. Instrument detection is a first step for augmenting the field of view with instrument description and characteristics. Instrument detection and tracking is a crucial part in studies of surgeon’s expertise [9] and in all domains where a human operator needs to practice a fine orchestra of micro-movements to achieve mastery in eye-hand coordination.

REFERENCES

- [1] E. Evgeniou, H. Walker, and S. Gujral, “The role of simulation in microsurgical training,” *Journal of Surgical Education*, 2017.
- [2] N. L. Crosby, J. B. Clapson, H. J. Buncke, and L. Newlin, “Advanced non-animal microsurgical exercises,” *Microsurgery*, vol. 16, no. 9, pp. 655–658, 1995.
- [3] I. Klein, U. Steger, W. Timmermann, A. Thiede, and H.-J. Gassel, “Microsurgical training course for clinicians and scientists at a german university hospital: A 10-year experience,” *Microsurgery*, vol. 23, no. 5, pp. 461–465, 2003.
- [4] H. Afkari, R. Bednarik, S. Mäkelä, and S. Eivazi, “Mechanisms for maintaining situation awareness in the micro-neurosurgical operating room,” *International Journal of Human-Computer Studies*, vol. 95, pp. 1–14, 2016.
- [5] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, “Face recognition: A convolutional neural-network approach,” *IEEE transactions on neural networks*, vol. 8, no. 1, pp. 98–113, 1997.
- [6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [7] S. Eivazi, R. Bednarik, M. Tukiainen, M. von und zu Fraunberg, V. Leinonen, and J. E. Jääskeläinen, “Gaze behaviour of expert and novice microneurosurgeons differs during observations of tumor removal recordings,” in *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM, 2012, pp. 377–380.
- [8] A. C. Furtado, I. Cheng, E. Fung, B. Zheng, and A. Basu, “Low resolution tool tracking for microsurgical training in a simulated environment,” in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 3843–3846.
- [9] X. Jiang, B. Zheng, and M. S. Atkins, “Video processing to locate the tooltip position in surgical eye-hand coordination tasks,” *Surgical innovation*, vol. 22, no. 3, pp. 285–293, 2015.
- [10] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [12] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, “A survey on deep learning in medical image analysis,” *Medical image analysis*, vol. 42, pp. 60–88, 2017.
- [13] F. Xing, Y. Xie, H. Su, F. Liu, and L. Yang, “Deep learning in microscopy image analysis: A survey,” *IEEE Transactions on Neural Networks and Learning Systems*, 2017.
- [14] C. Zhang, W. Tavanapong, J. Wong, P. C. de Groen, and J. Oh, “Real-time instrument scene detection in screening gi endoscopic procedures,” in *Computer-Based Medical Systems (CBMS), 2017 IEEE 30th International Symposium on*. IEEE, 2017, pp. 720–725.
- [15] A. Dutta, A. Gupta, and A. Zissermann, “VGG image annotator (VIA),” <http://www.robots.ox.ac.uk/vgg/software/via/>, 2016, accessed: 1.3.2018.
- [16] A. Martin, A. Ashish, B. Paul, B. Eugene, and Z. Chen, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, software available from tensorflow.org. [Online]. Available: <https://www.tensorflow.org/>
- [17] J. A. Hernesniemi, J. Choque-Velasquez, D. Kozyrev, P. Thiarawat *et al.*, “Hernesniemi’s 1001 and more microneurosurgical videos,” 2017.
- [18] S. Dodge and L. Karam, “Understanding how image quality affects deep neural networks,” in *Quality of Multimedia Experience (QoMEX), 2016 Eighth International Conference on*. IEEE, 2016, pp. 1–6.