

Neuro-Endo-Trainer-Online Assessment System (NET-OAS) for Neuro-Endoscopic Skills Training

Vinkle Kumar Srivastav*, Britty Baby*, Ramandeep Singh[†], Prem Kalra* and Ashish Suri[‡]

*Amar Nath and Shashi Khosla School of IT

Indian Institute of Technology Delhi, Hauz Khas, Email: vinkle.kumar@gmail.com

[†]Center for Biomedical Engineering

Indian Institute of Technology Delhi, Hauz Khas

[‡]Department of Neurosurgery,

All India Institute of Medical Sciences, New Delhi

Abstract—Neuro-endoscopy is a challenging minimally invasive neurosurgery that requires surgical skills to be acquired using training methods different from the existing apprenticeship model. There are various training systems developed for imparting fundamental technical skills in laparoscopy where as limited systems for neuro-endoscopy. Neuro-Endo-Trainer was a box-trainer developed for endo-nasal transsphenoidal surgical skills training with video based offline evaluation system. The objective of the current study was to develop a modified version (Neuro-Endo-Trainer-Online Assessment System (NET-OAS)) by providing a stand-alone system with online evaluation and real-time feedback. The validation study on a group of 15 novice participants shows the improvement in the technical skills for handling the neuro-endoscope and the tool while performing pick and place activity.

Index Terms—Neuro-endoscopy; Vision based surgical skills assessment; surgical skills training; Neuro endo trainer; online evaluation

I. INTRODUCTION

MINIMALLY invasive neurosurgical procedures have gained the popularity in recent years due to the reduction in postoperative recovery time, morbidity, hospitalization time and cost of patient care [1]. It provides the neurosurgeon with a better visualization method of the complex surgical site with reduced damage to the intricate anatomy of the brain. Neuro-endoscopy is a minimally invasive neurosurgical procedure that uses an endoscope image projected on the 2-dimensional display to access the interior deep structures. The margin of error is minimal and the existing apprenticeship based method of training is not suitable. It requires training for eye-hand coordination, depth perception, and bimanual dexterity. The simulation-based training outside the operating room is getting wide acceptance due to the provision of repeated practice, objective evaluation, real-time feedback and staged development of skills without the supervision of an expert surgeon [2].

Simulation-based training in neuro-endoscopy varies from low-fidelity natural simulations, box trainers, part-task trainers, to intermediate-fidelity synthetic simulators, virtual reality simulators and high-fidelity cadavers and animal models. The box-trainers or part-task trainers are designed to impart training for fundamental technical skills of instrument handling and

eye-hand coordination. The synthetic simulators and virtual reality trainers provide training for anatomy and procedures but give limited haptic feedback. The high-fidelity simulations on cadavers and animals provide training for anatomy and procedures along with haptic feedback and realism [3]–[7].

The evaluation of the surgical activity on the various simulation systems is platform-specific. The assessment methods can be based on direct observation, error metric of the task, sensor-based evaluation of the motion and video-based evaluation of the activity or combination of these. The validation studies on Neurosurgery Education and Training School-Skills Assessment Scale (NETS-SAS) identifies the independent parameters of neurosurgery skills as hand-eye coordination, instrument-tissue manipulation, dexterity, flow of procedure and effectualness [8]. These parameters can be analyzed by the video-based evaluation systems that monitor the activity and movement of the surgeon's hands or tools. The video recording of the activity also provides an opportunity to validate the evaluation using subjective methods.

The video based automatic assessment system can be of two types; offline evaluation and online evaluation. Offline evaluation systems acquire the activity video at reasonable rate and stores the video stream for further analysis. The online evaluation system uses the frame-by-frame analysis, that simultaneously evaluate the activity and also stores it for future reference.

Neuro-Endo-Trainer was a box trainer developed for providing skills training for endo-nasal transsphenoidal surgery (ENTS). It was a pick-and-place task trainer that provides the training for basic fundamental skills using standard variable angled neuro-endoscopes [8]. The evaluation method includes video-based offline evaluation using an auxiliary camera mounted at the top of the box [9]. The existing method of training on Neuro-Endo-Trainer involves the pick and place of one of the six rings in a predefined pattern under the assistance of technical personnel. The activity performed is sub-divided into sub-activity based on the state of the tool and the rings. The sub-activity can be “stationary”, “picking” or “moving”. The state machine is determined using video processing that includes the tooltip tracking, background segmentation, and ring segmentation. The definition of state machine with

the heuristics determined from the video, causes uncertainty and requires a robust task definition system. Therefore, the hardware of the Neuro-Endo-Trainer was augmented with automatic LED-based task definition to determine the state machine. We have developed a stand-alone training system with Neuro-Endo-Trainer to provide online assessment and real-time feedback and defined it as Neuro-Endo-Trainer-Online Assessment System (NET-OAS). Our online automatic assessment system analyzes the activity frame-by-frame and categorizes it as a sub-activity. The relevant parameters of skills training are identified by statistical analysis of the sub-activity. It provides a warning to the trainee neurosurgeon when they make mistakes and provide a detailed synopsis at the end of the activity. The aim of the current study is to validate the developed NET-OAS to establish the level of skills acquisition after staged practice.

II. BACKGROUND

The low fidelity box-trainers are widely available for laparoscopic skills training [10], [11] whereas they are limited for neuro-endoscopy. The evaluation system for these trainers can be based on subjective or objective measures. The objective evaluation includes Likert-scale based direct observation, sensor-based evaluation and computerized video analysis. The webcam based endoscopic endonasal trainer developed by Hirayama et al. studied the effectualness of the training by evaluating the performance on LapSim simulator before and after the training [3]. Neuro-Endo-Trainer SkullBase-Task-GraspPickPlace developed by Raman et.al was validated using subjective evaluation on different target groups [8].

The video-based evaluation of the surgical activity includes the tracking of the tooltip or tracking the surgeon's hands. There are evaluation systems that use statistical color based image segmentation and tool tracking to identify the tool position and orientation [12], [13]. The automated skills evaluation method in minimally invasive laparoscopic surgeries were done by segmenting the task into sub-tasks (Therbligs) and their kinematic analysis [14]. The feature based tool tracking combined with region-based level set segmentation was used to obtain 3D pose estimation of the instruments and to evaluate the psychomotor skills [15]. There are methods that capture the activity of the subject and track the hand movements using multiple camera feeds [16]. Neuro-Endo-Activity-Tracker provided a video-based automatic evaluation using Gaussian Mixture based background subtraction and tracking of the tooltip using Tracking-Learning-Detection algorithm [9].

III. METHODOLOGY

NET-OAS consists of low-cost endoscopic system of USB based endoscopic camera that captures the video at 25 fps, variable-angled scopes (0° , 30° , 45°), LED-based light source, Neuro-Endo-Trainer SkullBase-Task-GraspPickPlace box-trainer mounted with GigE based auxiliary camera, and online evaluation software.

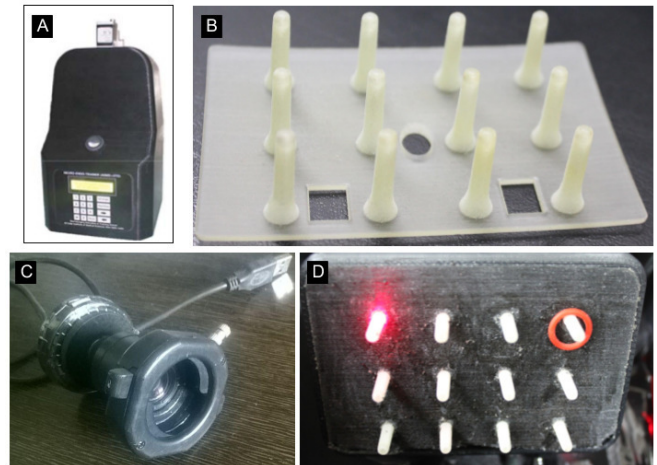


Fig. 1. A. Neuro-Endo-Trainer SkullBase-Task-GraspPickPlace box-trainer mounted with GigE based auxiliary camera, B. Transparent front-part of the peg plate, C. USB camera with endoscope coupler, D. Peg plate with LED

A. NET-OAS hardware design

The online evaluation system consists of a LED-based task indication method which helps the user to place the ring on the illuminated peg without the assistance of any technician. The peg was illuminated to provide the indication for placement of the ring. The peg plate was printed in two parts: front part of the peg was printed using transparent material by Stereolithography (SLA) technique and back part of the plate was printed using fused deposition modeling (FDM) technique and then both parts were joined using a strong adhesive. The LED array was connected to control circuit using a multiplexer (CD74HC4067). The control circuit consists of ATMEGA328 8 bit micro-controller for the processing, MCP23017 I/O port expander for I/O expansion, 16x2 LCD for display, keypad to provide input, servo motor to control the peg plate and FT232RL serial communication chip to communicate with the PC using serial communication protocol. There are two cameras in the setup; Low-cost USB based endoscopic camera for the visualization of the site that captures feed at 25fps and GigE based auxiliary camera (Basler ACE) capturing at 50 fps for the online evaluation and real-time feedback. The hardware components of NET-OAS is shown in Fig. 1.

B. NET-OAS software design

The software system of NET-OAS uses a multi-threaded program that processes the two camera streams independently, which maintains the real-time requirement of the system. The complete flow diagram of the NET-OAS is shown in Fig.2 and its user interface is shown in Fig.3. It shows endoscopic and auxiliary streams, options to add the user to the database, configure serial port parameters, select the level of training and option to perform calibration if required. When the user hit the Run button, a new window opens the endoscopic stream with screen display of real-time feedback. After the completion of the activity, the results are shown to the user.

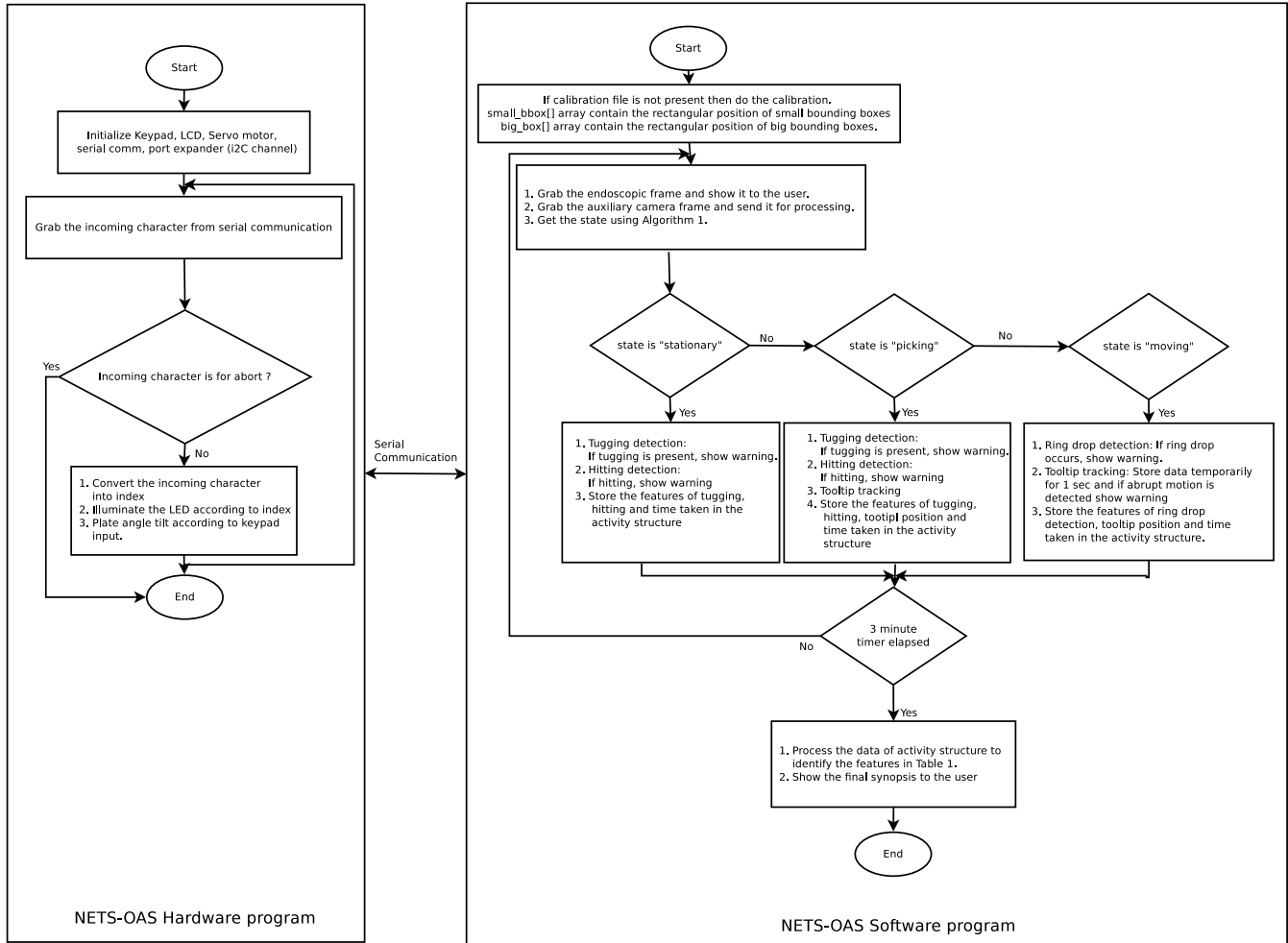


Fig. 2. Flow Diagram of NET-OAS

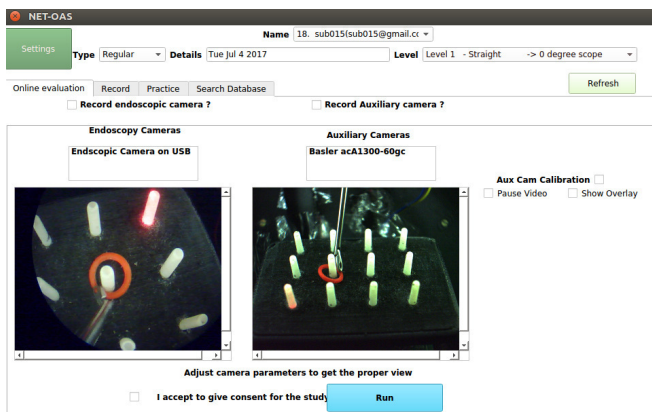


Fig. 3. User interface of NET-OAS

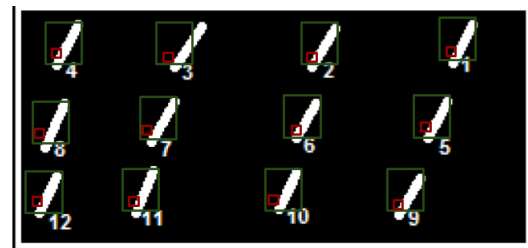


Fig. 4. Bounding box of peps

The main components of the software system are as follows:

1) *Calibration setup:* One-time calibration involves peg-segmentation, ring segmentation, and tooltip bounding box

selection and storing the parameters in the calibration file. The `small_bbox[]` contains the rectangular location of small bounding boxes, `big_bbox[]` contains the location of big bounding boxes as shown in Fig.4. These arrays are used to determine the state machine explained in Algorithm 1. When the software starts, it loads the parameters from the calibration file otherwise prompt the user to perform the calibration.

2) *State machine estimation:* The activity on the NET-OAS in a particular frame can be any of the following

Algorithm 1 Determine the state machine

```

Read the calibration file; Initialize small_bbox[], big_bbox[],
thresh_stationary, thresh_picking, thresh_moving
old_id  $\leftarrow$  -1; current_id  $\leftarrow$  -1;
index_set  $\leftarrow$  true; status  $\leftarrow$  "stationary";
function GET-STATE(image)
  if index_set then
    index_set  $\leftarrow$  false
    for k = 0; k < 12; k++ do
      seg_image  $\leftarrow$  ringSegmentation(image);
      sum_pixels  $\leftarrow$  seg_image[small_bbox[k]];
      if sum_pixels  $\leq$  thresh_picking then
        old_index  $\leftarrow$  k + 1;
        break;
      end if
    end for
    current_index  $\leftarrow$  random(1 – 12);
    litLED(current_index)
    seg_image  $\leftarrow$  ringSegmentation(image);
    s_old_small  $\leftarrow$  seg_image[small_bbox[old_id]];
    s_old_big  $\leftarrow$  seg_image[small_bbox[old_id]];
    s_current_small  $\leftarrow$  seg_image[small_bbox[current_id]];
    if status == "stationary" then
      if s_old_small  $\geq$  thresh_stationary then
        status  $\leftarrow$  "stationary";
      else
        status  $\leftarrow$  "picking";
      end if
    else if status == "picking" then
      if s_old_big  $\geq$  thresh_picking then
        status  $\leftarrow$  "picking";
      else
        status  $\leftarrow$  "moving";
      end if
    else if status == "moving" then
      if s_current_small  $\geq$  thresh_moving then
        status  $\leftarrow$  "stationary";
        old_id  $\leftarrow$  current_id;
        current_id  $\leftarrow$  random(1 – 12);
        litLED(current_id)
      else
        status  $\leftarrow$  "moving";
      end if
    end if
  return status
end function

```

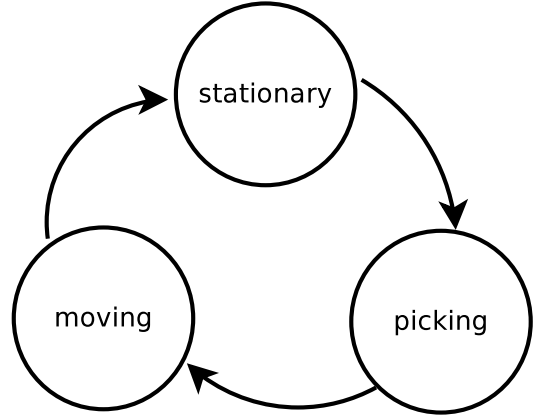


Fig. 5. State-machine

sub-activity: "stationary", "picking" or "moving". The state machine is initialized with the "stationary" state and the states are updated according to the movement of the ring. The "stationary" state is defined when the ring is stationary and the tool is present/absent. The "picking" state is defined when the tool is near the peg trying to grab the ring till the ring moves out of the peg. The "moving" state is defined when the ring has moved out of the peg until it is placed on the illuminated destination peg. Once the ring has been placed on the peg, the ring segmentation output in the bounding box changes and another peg is illuminated randomly. The state machine is unidirectional and cyclic as shown in Fig.5. The algorithm for state machine estimation is explained in Algorithm 1. Function *ringSegmentation(image)* perform the ring segmentation on the input frame and *litLED(int number)* function illuminate the corresponding peg given in its argument.

3) *Tracking Algorithm*: Tracking-Learning-Detection (TLD) algorithm is used to track the tooltip. TLD initializes from the bounding box and tracking model, retrieved from the calibration file. It is a robust tracking algorithm which tracks the tooltip under blurred conditions and various transformations. The tracking is based on median flow tracker which track the tooltip frame-to-frame and measure the tracking error using efficiency of backtracking. The detection thread is a 3-stage sliding window cascaded classifier, which consists of variance filter, random forest, and nearest neighbor classifier. At the end of the 3rd stage, it provides a set of windows that localizes the appearance of the tool tip. It predicts the next location of the tool tip having the minimum error in tracking or detection stage. The remaining set of appearances is fed to the negative class for better generalization of the tool tip model. Tracking of the tool using TLD algorithm is shown in Fig.6 A. [17].

4) *Ring Drop Detection*: The dropping of the ring is determined in the "moving" state if distance between the tool tip bounding box (determined by TLD) and the *ringSegmentation(image)* is more than a predefined threshold. Fig.6 B shows the image of the ring drop condition.

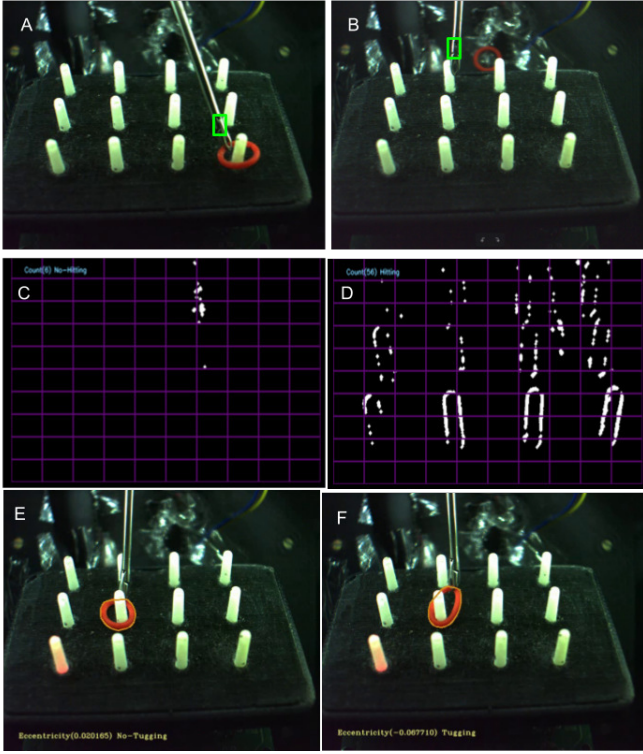


Fig. 6. Auxiliary camera frame analysis showing: A. Tracking of the tool using TLD algorithm, B. Ring drop determined by the distance between tool-tip and ring segmentation. C. No Hitting D. Hitting determined by counting the subwindows having significant number of contours, E. No- Tugging F. Tugging determined by eccentricity analysis of the ring contour

5) *Hitting Detection*: The hitting of the peg board happens due to poor depth perception of the user. The hitting is detected using image analysis of the successive frames. The difference image is divided into 10x10 grids and hitting is recorded by identifying the number of grids that shows significant movement. The hitting threshold is set experimentally and the Fig.6 C shows the case of no hitting and Fig.6 D shows a hitting instance output.

6) *Tugging detection*: The tugging is detected by analyzing the deformation of the ring in the “stationary” and “picking” state. The ring is segmented based on the hue value obtained from the calibration file. Due to the overlapping of the tool or peg, $ringSegmentation(image)$ results in two or more contours. The contour with maximum size and the nearest contours are determined and combined. The

$$eccentricity = \frac{\mu_{2,0} + \mu_{0,2} + \sqrt{(\mu_{2,0} - \mu_{0,2})^2 + 4(\mu_{1,1})^2}}{\mu_{2,0} + \mu_{0,2} - \sqrt{(\mu_{2,0} - \mu_{0,2})^2 + 4(\mu_{1,1})^2}}$$

value of the combined contour is sufficient to determine the deformation of the ring in case of tugging. The eccentricity threshold corresponding to tugging is set experimentally.

7) *Tracking data analysis*: Tracking data analysis is done to identify motion smoothness and sudden jerk of the tool tip motion in the “moving” state. Smoothness of the path is measured by taking the standard deviation of the first

TABLE I
SELECTED FEATURES FOR NET-OAS

Measure from NETS-SAS	Selected objective measure for NET-OAS
<i>Grasping</i>	Average time taken to grasp
	Number of tugging events
<i>Eye-hand coordination</i>	Number of hitting events
	Intensity with which hitting happened
<i>Dexterity</i>	Time taken for moving ring from one peg to another
	Average number of moves
	Smoothness of the path
	Arc length of the path
<i>Instrument tissue manipulation</i>	Number of times curvature value exceeded threshold
<i>Effectualness</i>	Number of times ring dropped

derivative of the tracking data, Arc length of the path is measured by counting number of pixels of the tracking data in the “moving” state. Curvature at each point of tracking data is computed using

$$\kappa = \frac{|\left(\frac{\partial x}{\partial t} * \frac{\partial^2 y}{\partial t^2}\right) - \left(\frac{\partial y}{\partial t} * \frac{\partial^2 x}{\partial t^2}\right)|}{\left(\frac{\partial x}{\partial t}^2 + \frac{\partial y}{\partial t}^2\right)^{\frac{3}{2}}}$$

8) *Real time feedback*: At each frame, the algorithm identifies the current state and provide real time feedback for hitting, tugging and ring drop. Motion smoothness feedback is provided after processing frames of last 1 second. The output is displayed on the endoscopic screen to warn the user. This helps the user to learn and correct the mistakes accordingly.

9) *Feature Extraction and final synopsis*: The activity data structure stores the current sub-activity (“stationary”, “picking” or “moving”) and its related parameters as shown in Table 1. At the end of the activity, the data is processed to give the final synopsis to the user.

IV. EXPERIMENTATION AND RESULTS

A group of 15 novices participated in the study of validation of NET-OAS, who were students from a technical university without any medical training. The demo video demonstrating the good and bad endoscopy practice on Neuro-Endo-Trainer was shown before the practice session. There was a pre-test followed by two sessions and a post-test. The pre-test and post-test included the most difficult task level of 45° scope with right tilt plate. Each activity was programmed to be of 3 minutes duration. The first session consisted of practice using 0° and 30° scopes and with straight, left and right tilts of the plate. The second session was conducted three days later and consisted of practice using 30° and 45° scopes and with straight, left and right tilts of the plate. Fig. 7 shows the graph of objective measure for NET-OAS w.r.t training session. The noticeable changes were the increased

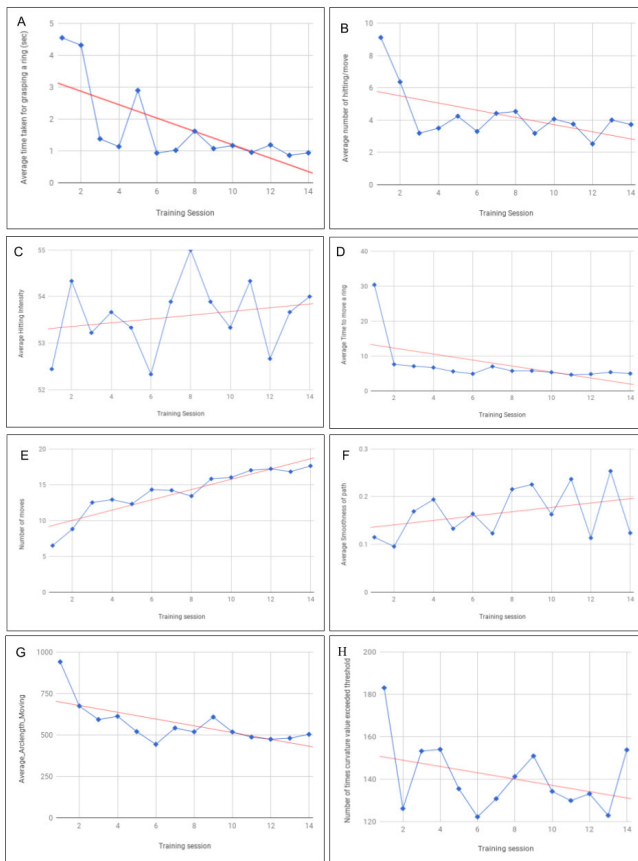


Fig. 7. Validation study results: Horizontal axes is the training session, blue marker shows the data point and red line shows the trend-line: A. Average time of grasping the ring, B. Average number of hitting, C. Average hitting intensity D. Average time to move a ring, E. Total number of rings placed F. Average smoothness of the tool tip in “moving” state, G. Average Arc length of the tool tip in “moving” state, H. Number of times curvature exceeded the threshold value or sudden jerk.

average number of moves and average smoothness of the path. There were decreased number and hitting instances, grasping time, average arc length and sudden jerk motion. The self-assessment feedback obtained from the user also shows that the training session on the NET-OAS made them acquainted with the system.

1) *Machine learning for validation study:* For the validation study, activity data obtained from 15 novices (pre-test, post-test, 1st trial of session 1 and last trial of session 2) was considered. Pre-test data was considered as ‘class novice’ and post-test data was considered as ‘class-improved’. The SVM classifier was trained with 11-dimensional feature vector of these classes. For testing, 1st trial of session 1 was considered as ‘class novice’ and the last trial of session 2 was considered as ‘class improved’. The SVM classifier on the testing data classifies feature set of the 1st trial as ‘class novice’ and the last trial of session 2 as ‘class improved’ with the accuracy of 88%.

The practice session example on the NET-OAS and the real-

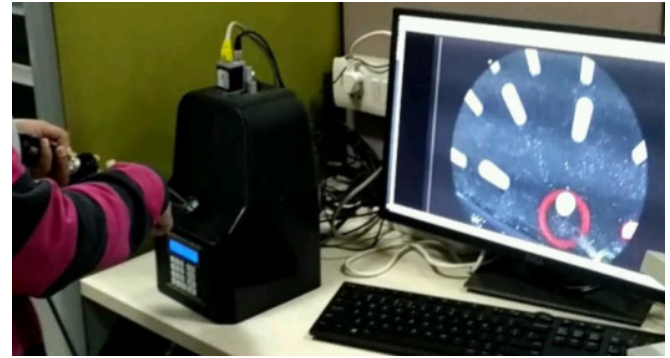


Fig. 8. Training on the NET-OAS

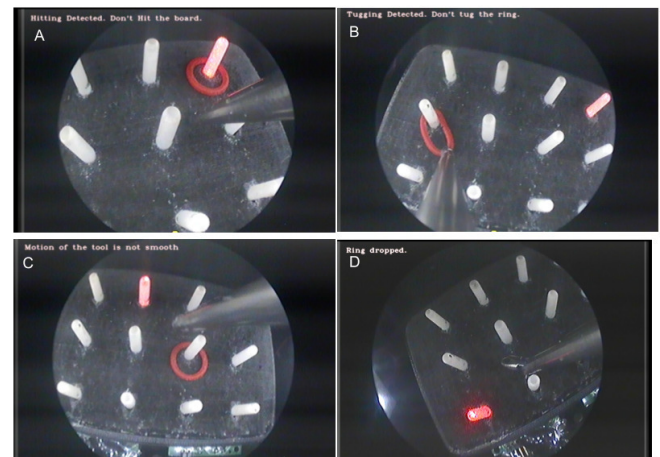


Fig. 9. Real-time feedback to trainee A) Hitting B) Tugging C) Motion smoothness D) Ring Drop

time feedback provided to the trainee while performing the activity is as shown in Fig. 8 and Fig. 9 respectively.

V. DISCUSSION

The improvements of NET-OAS as compared to the earlier version include: a complete standalone system, automatic task definition using LED array and serial communication with the hardware, tugging detection algorithm, and ring drop detection. The study used the auxiliary camera for the evaluation of the activity and has not used the endoscopic feed for evaluation.

The main objective of the study was to validate the NET-OAS on completely novice participants to identify whether there is any improvement in skills acquisition. The results show that after stipulated training on the NET-OAS, the participant improved his/her skills on manipulating the endoscope and tool irrespective of their background. The study can be extended to the intermediate trainee neurosurgeons and experts.

ACKNOWLEDGMENT

We would like to thank all the participants, research scholars of Indian Institute of Technology Delhi who took part

in the study, and the team of Neurosurgery Education and Training School for their support. This work is supported by Department of Health Research, Ministry of Health and Family Welfare, Govt. of India Project Code No: GIA/3/2014-DHR, Department of Science and Technology (DST), Ministry of Science and Technology, Govt. of India Project Code No: SR/FST/LSII-029/2012.

REFERENCES

- [1] R. Abbott, "History of neuroendoscopy," *Neurosurgery Clinics of North America*, vol. 15, no. 1, pp. 1–7, 2004.
- [2] M. Bridges and D. L. Diamond, "The financial impact of teaching surgical residents in the operating room," *The American Journal of Surgery*, vol. 177, no. 1, pp. 28–32, 1999.
- [3] R. Hirayama, Y. Fujimoto, M. Umegaki, N. Kagawa, M. Kinoshita, N. Hashimoto, and T. Yoshimine, "Training to acquire psychomotor skills for endoscopic endonasal surgery using a personal webcam trainer: Clinical article," *Journal of neurosurgery*, vol. 118, no. 5, pp. 1120–1126, 2013.
- [4] R. Singh, V. K. Srivastav, B. Baby, N. Damodaran, and A. Suri, "A novel electro-mechanical neuro-endoscopic box trainer," in *Industrial Instrumentation and Control (ICIC), 2015 International Conference on*. IEEE, 2015, pp. 917–921.
- [5] G. Rosseau, J. Bailes, R. del Maestro, A. Cabral, N. Choudhury, O. Comas, P. Debergue, G. De Luca, J. Hovdebo, D. Jiang *et al.*, "The development of a virtual simulator for training neurosurgeons to perform and perfect endoscopic endonasal transsphenoidal surgery," *Neurosurgery*, vol. 73, pp. S85–S93, 2013.
- [6] S. Wolfsberger, M.-T. Forster, M. Donat, A. Neubauer, K. Bühler, R. Wegenkittl, T. Czech, J. A. Hainfellner, and E. Knosp, "Virtual endoscopy is a useful device for training and preoperative planning of transsphenoidal endoscopic pituitary surgery," *min-Minimally Invasive Neurosurgery*, vol. 47, no. 04, pp. 214–220, 2004.
- [7] J. Fernandez-Miranda, J. Barges-Coll, D. Prevedello, J. Engh, C. Snyderman, R. Carrau, P. Gardner, and A. Kassam, "Animal model for endoscopic neurosurgical training: technical note," *min-Minimally Invasive Neurosurgery*, vol. 53, no. 05/06, pp. 286–289, 2010.
- [8] R. Singh, B. Baby, N. Damodaran, V. Srivastav, A. Suri, S. Banerjee, S. Kumar, P. Kalra, S. Prasad, K. Paul *et al.*, "Design and validation of an open-source, partial task trainer for endonasal neuro-endoscopic skills development: Indian experience," *World neurosurgery*, vol. 86, pp. 259–269, 2016.
- [9] B. Baby, V. K. Srivastav, R. Singh, A. Suri, and S. Banerjee, "Neuro-endo-activity-tracker: An automatic activity detection application for neuro-endo-trainer: Neuro-endo-activity-tracker," in *Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on*. IEEE, 2016, pp. 987–993.
- [10] J. Dankelman, M. Chmarra, E. Verdaasdonk, L. Stassen, and C. Grimbbergen, "Fundamental aspects of learning minimally invasive surgical skills," *Minimally Invasive Therapy & Allied Technologies*, vol. 14, no. 4-5, pp. 247–256, 2005.
- [11] M. K. Chmarra, N. H. Bakker, C. A. Grimbbergen, and J. Dankelman, "Trendo, a device for tracking minimally invasive surgical instruments in training setups," *Sensors and Actuators A: Physical*, vol. 126, no. 2, pp. 328–334, 2006.
- [12] C. L. Y. W. D. Uecker and Y. Wang, "Image analysis for automated tracking in robot-assisted endoscopic surgery," in *Proc. 12th Int'l Conf. Pattern Recognition*, 1994, pp. 88–92.
- [13] G.-Q. Wei, K. Arbter, and G. Hirzinger, "Real-time visual servoing for laparoscopic surgery. controlling robot motion with color image segmentation," *IEEE Engineering in Medicine and Biology Magazine*, vol. 16, no. 1, pp. 40–45, 1997.
- [14] S.-K. Jun, M. S. Narayanan, P. Agarwal, A. Eddib, P. Singhal, S. Garimella, and V. Krovi, "Robotic minimally invasive surgical skill assessment based on automated video-analysis motion studies," in *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on*. IEEE, 2012, pp. 25–31.
- [15] M. Allan, S. Thompson, M. J. Clarkson, S. Ourselin, D. J. Hawkes, J. Kelly, and D. Stoyanov, "2d-3d pose tracking of rigid instruments in minimally invasive surgery," in *International Conference on Information Processing in Computer-assisted Interventions*. Springer, 2014, pp. 1–10.
- [16] Q. Zhang, L. Chen, Q. Tian, and B. Li, "Video-based analysis of motion skills in simulation-based surgical training," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2013, pp. 86 670A–86 670A.
- [17] Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-detection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 7, pp. 1409–1422, 2012.