OPEN ACCESS
Journal of Medical Robotics Research, Vol. 6, Nos. 3 & 4 (2021) 2150004 (10 pages)

DOI: 10.1142/S2424905X21500045

Journal of Medical Robotics Research
https://www.worldscientific.com/worldscinet/jmrr

Mapping Surgeons Hand/Finger Movements to Surgical Tool Motion During Conventional Microsurgery Using Machine Learning

Mohammad Fattahi Sani*, Raimondo Ascione†, Sanja Dogramadzi‡,§

*Bristol Robotics Laboratory, University of the West of England, Bristol, UK

†Bristol Heart Institute and Translational Biomedical Research Centre, Faculty of Health Science, University of Bristol, Bristol, UK

‡Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffiled, UK

Purpose: Recent developments in robotics and artificial intelligence (AI) have led to significant advances in healthcare technologies enhancing robot-assisted minimally invasive surgery (RAMIS) in some surgical specialties. However, current human–robot interfaces lack intuitive teleoperation and cannot mimic surgeon's hand/finger sensing required for fine motion micro-surgeries. These limitations make teleoperated robotic surgery not less suitable for, e.g. cardiac surgery and it can be difficult to learn for established surgeons. We report a pilot study showing an intuitive way of recording and mapping surgeon's gross hand motion and the fine synergic motion during cardiac micro-surgery as a way to enhance future intuitive teleoperation.

Methods: We set to develop a prototype system able to train a Deep Neural Network (DNN) by mapping wrist, hand and surgical tool real-time data acquisition (RTDA) inputs during mock-up heart micro-surgery procedures. The trained network was used to estimate the tools poses from refined hand joint angles. Outputs of the network were surgical tool orientation and jaw angle acquired by an optical motion capture system.

Results: Based on surgeon's feedback during mock micro-surgery, the developed wearable system with light-weight sensors for motion tracking did not interfere with the surgery and instrument handling. The wearable motion tracking system used 12 finger/thumb/wrist joint angle sensors to generate meaningful datasets representing inputs of the DNN network with new hand joint angles added as necessary based on comparing the estimated tool poses against measured tool pose. The DNN architecture was optimized for the highest estimation accuracy and the ability to determine the tool pose with the least mean squared error. This novel approach showed that the surgical instrument's pose, an essential requirement for teleoperation, can be accurately estimated from recorded surgeon's hand/finger movements with a mean squared error (MSE) less than 0.3%.

Conclusion: We have developed a system to capture fine movements of the surgeon's hand during micro-surgery that could enhance future remote teleoperation of similar surgical tools during micro-surgery. More work is needed to refine this approach and confirm its potential role in teleoperation.

Keywords: Robot-assisted surgery; minimally invasive surgery; machine learning; hand tracking; real-time low-cost hand tracking; feature extraction.

JMRR

Received 29 March 2021; Revised 20 July 2021; Accepted 11 August 2021; Published 18 October 2021. This paper was recommended for publication in its revised form by editorial board member Dan Stoianovici.
Email Addresses: §s.dogramadzi@sheeld.ac.uk

1. Introduction

Robot-Assisted Surgery (RAS) is preferred to conventional surgery in several clinical applications including, e.g. urology, cardiovascular, due to reduced invasiveness, superior ergonomics, precision, dexterity and intuitive interaction [1], resulting at times in shorter procedure and hospitalization times [2]. Intuitiveness of surgical teleoperation in RAS is essential in ensuring safety while obtaining the right level of procedural accuracy and effectiveness. Effective teleoperation depends on accurate mapping of the surgeon's hand/fingers operating motion and the flawless translation of these fine movements from the surgeon's master to the surgical instrument slave.

Sub-optimal teleoperation has limited the widespread use of safe and effective RAS. For example, while RAS is widely utilized in urology, its adoption is limited or nonexistent in micro-surgical specialties. Another limiting factor is that the surgeon's master of current RAS systems is very different from conventional microsurgical instruments which makes it difficult for established surgeons to adopt RAS unless subjecting themselves to a new training. Moreover, the precision of slave positioning in, e.g. Da Vinci surgical robot is not sufficient for surgeries requiring higher precision.

Cardiac surgery and other specialty areas have seen little RAS penetration and some safety concerns [1,3]. In cardiac surgery, for example, the limited access to the heart, no space is available for the slave instrument maneuvering and close proximity to other vital structures requires finer slave instrument movements and a superior teleoperation to that available in current robotic systems. For example, the Da Vinci master console uses a pair of handles [4] to control the slave end-effector with 3 degrees of freedom (DOF), open/close grasper and 2 DOF in the wrist. While this appears to be effective for urology, for MIS tools that require higher complexity and dexterity, the master side of the system should be refined to support the feasibility, safety and efficacy of micro-surgery specialties through smoother and more intuitive teleoperation.

Proposed methods to teleoperate a surgical robot range from using handles at the surgeon's master console [4] to a touch screen control [5] supplemented with gaze, voice and foot pedals control [6]. Commonly used master stations in robotic surgery such as the Da Vinci master station, Phantom Omni, haptic device neuroArm system [7] were compared with novel approaches such as wireless or wearable data gloves and upper-body exoskeleton masters, suggesting that future RAS technologies are more likely to use wearables as master device [4].

Hand/finger tracking has been investigated for applications ranging from teleoperation to motion analysis. To this end, inertial measurement units (IMUs), optical sensors, exoskeletons, magnetic sensing, and flex sensors-based systems have been used [8]. IMUs have also been used to complement or compensate errors of Kinect depth camera for skeletal tracking [9]. However, depth camera is vulnerable to occlusions which prevent reliable pose estimation [10]. IMU sensors on thumb and index finger are sufficiently precise to authenticate in-air-handwriting signature using a support vector machine (SVM) classifier [11].

The concept of Robotic Learning from Demonstration (LfD) is attractive for teaching robot repetitive tasks from expert surgeon's demonstrations as shown in [12] for RAMIS or using surgical instruments trajectories to extract key features during complex surgical tasks [13]. Machine learning classification methods to evaluate surgeon's skills by processing data extracted from the Da Vinci system have been shown in [14] or by tracking MIS tools in [15]. While in most studies, surgeon's movements have been used to analyze generated trajectories, to the best of our knowledge, the actual dexterity of surgeon's hands/fingers during complex surgeries has not been studied [16].

Control of robotic minimally invasive surgical system using index finger and thumb gestures has been shown in [17,18]. However, using gestures to teleoperate a surgical robot is not intuitive or precise. In our earlier studies, we designed a hand tracking wearable system using IMU sensors placed on the hand digits and wrist to control a 4 DOF da Da Vinci Endowrist instrument [19]. For this study, we implemented our 12 DOF tracking wearable device to capture and analyze complex motion of surgeon's hand during cardiac surgery procedures which helped us define the digits range of motion, workspace and the motion rate of change [8].

Here, we show how surgeon's hand and wrist motion during open access micro-surgery is mapped to the surgical instrument motion with the aim to enhance a more effective, fine resolution, multi-DOF teleoperation of surgical instruments for micro-surgery to be used in future surgical robotic systems. To this end, our ultimate goal was to develop the first anthropomorphic prototype master concept based on learning the complexity of the fine synergic micro-surgical motion.

Our results demonstrate how surgeon's hand and wrist motion during open access micro-surgery can be mapped to the surgical instrument motion with the aim to enhance a more effective, fine resolution, multi-DOF teleoperation of surgical instruments for micro-surgery to be used in future surgical robotic systems.

Our ultimate contribution presented in this paper is a demonstration of the feasibility of using machine learning to map the anthropomorphic master concept to the fine synergic micro-surgical instrument motion with multiple DOFs.

2. Methodology

To develop the anthropomorphic prototype master concept based on mapping surgeon's hands/fingers to the

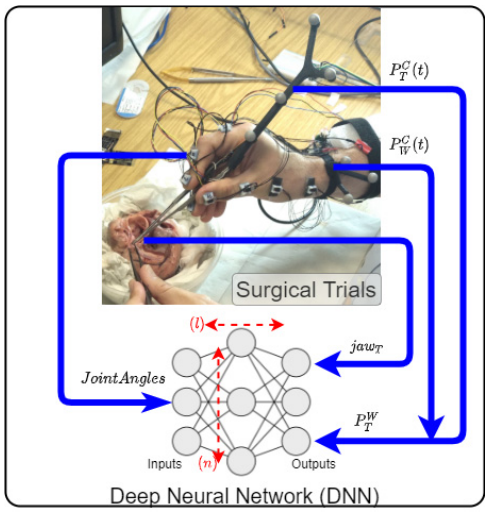


Fig. 1. The proposed concept: train a network during the open surgery and using that to map the hand finger motions to the surgical instrument. **Network outputs:** Four values. P_T^C: tool pose in camera frame, P_W^C: wrist pose in the camera frame. They will make P_T^W as tool pose with respect to the wrist (). jaw_T: tool jaw angle (opening closing of the tool). **Network inputs:** 15 joints in Euler angles. * l and n represent number of hidden layers and number of neurons in each hidden layer, respectively.

fine synergic motions of surgical instruments, we disregard the gross motion of the robotic shaft holding the surgical instrument. The approach used is illustrated in Fig. 1.

A light weight wearable hand/finger/tool-status (open/close) custom made tracking system based on clinical feedback was designed to allow natural movements during surgery and capture complex movements generated during mock cardiac surgery scenarios. An optimized ANN architecture based on deep learning was built to estimate surgical tool position and orientation and validate the proposed hand-tool relationship and utilization of each hand/finger joint while performing the tool motion in a specific surgical task.

First, we developed motion tracking systems of the hand and the selected surgical tools — fine Castroviejos needle-holder and surgical forceps operated through the typical three-finger approach thumb, index and middle finger. The tracking system captured two sets of data during typical mock cardiac surgery procedure performed by a senior heart surgeon. We derived a relationship between the two consecutive experiments conducted *ex-vivo* on animal samples using Castroviejos surgical tools. Data capturing focused on surgeon's hand/wrist movements and concomitant Castroviejo tools poses. A trained neural network was used to map the Castroviejo motion using the hand joint angles. In addition, we used Gini metrics of decision tree regressor learning approach [20] to assess contribution of each finger and thumb joints in the performed surgical tasks.

2.1. *Hand and tool pose tracking*

Surgical operations require complex dexterous manipulations of both hands so merely tracking the digit joint angles is not sufficient to fully represent hand poses and map them to the tool. After initial observations and analysis of cardiac open surgical procedures, the following key motion aspects were targeted to be captured:

(1) Surgeon's hand joint angles (digit and wrist joints).
(2) Global position and orientation of surgeon's hands.
(3) Global position and orientation of the surgical instrument.
(4) Surgical Instrument's Jaw angle (opening/closing (O/C) of the surgical instrument).

We used a chain of 12 IMUs to track hand and wrist joint angles (Fig. 2(a)). Hand digits are comprised of three joints, metacarpophalangeal (MCP), proximal interphalangeal (PIP), and distal interphalangeal (DIP) joint and represented in the hand kinematic model shown in our previous work [8]. Thumb joints are Carpometacarpal (CMC), Metacarpophalangeal (MP), Interphalangeal (IP). Our hand/wrist tracking system is comparable with the latest commercial data gloves (e.g. VR Manus) but features more sensing points in order to account for fine motion and a larger number of DOFs. Therefore, the index and middle finger and the thumb, with the key roles in fine grasping and manipulation of Castroviejo instruments, are each tracked by 3 IMUs (Fig. 2(a)). Top side of the forearm (as a reference point) and top side of the palm are each tracked by separate IMUs in order to calculate wrist joints. The remaining IMU is placed on the side of the palm parallel to the IMU attached on the thumb metacarpal in order to better track the CMC joint. The Castroviejo instrument is more complex for tracking due to its shape and size, particularly its open/close function for which we used four strain gauges attached to each side of the Castroviejo instrument (Fig. 2(b)) in a full Wheatstone bridge configuration.

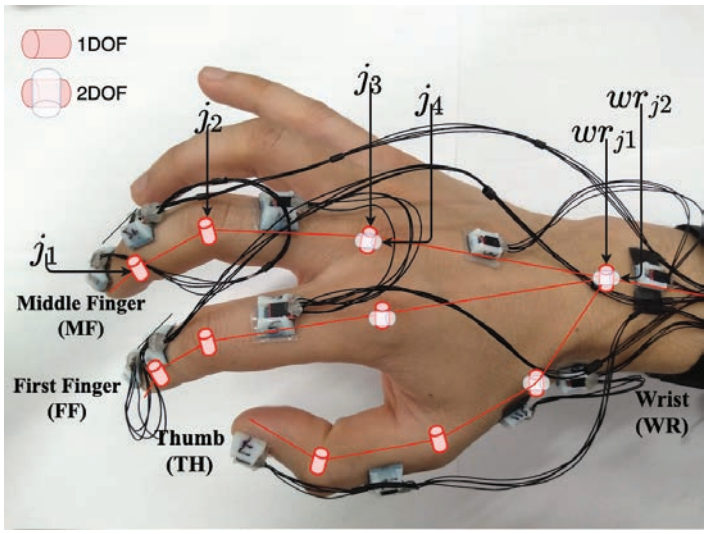
The global position and orientation of surgeon's hand and the surgical instrument are recorded using a Polaris Spectra optical sensor from Northern Digital Inc. (NDI) [21] and a set of holders for reflective infrared markers designed and attached to the wrist and the Castroviejo tool (Fig. 3 (right)). Polaris Spectra track the markers in 6 DOF and acquire the hand pose data using the proprietary NDI track software SawNDITracker library from Computer-Integrated Surgical Systems and Technology (CISST) [22] publishes data sent by NDI in the Polaris camera frame.

P_T^C and P_W^C are poses of the castro tool and the surgeon's wrist in the camera frame, respectively. We then calculate P_T as a tool pose with respect to the wrist:

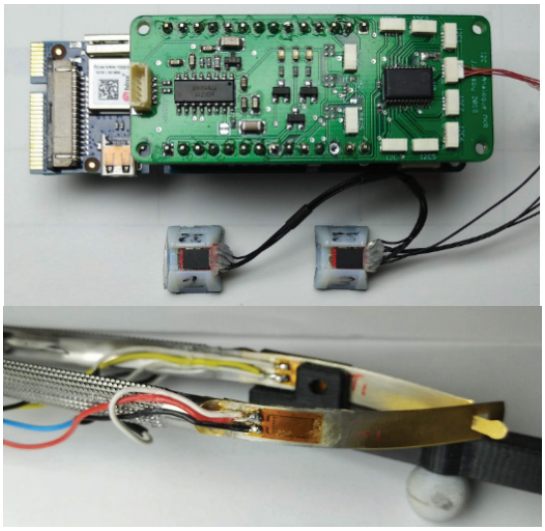
$P_T = [\phi_T(t), \theta_T(t), \psi_T(t)]$,

where $\phi_T(t)$, $\theta_T(t)$, and $\psi_T(t)$ are roll, pitch and yaw angles of Castroveijho tool in space.

The system architecture is presented in Fig. 4. The data are collected by: (1) IMU sensors, (2) strain gauges, and (3) Polaris motion capture system. We used a



(a) 12 IMU sensor attachments on the hand with joint names. (Note the single and double axis joint tracking. In the Middle Finger and the First Finger, j_1 and j_2 represent DIP and PIP joints, respectively, whereas j_3 and j_4 are two axes of the MCP joint. In the Thumb, j_1 and j_2 represent IP and MCP joints, respectively, whereas j_3 and j_4 are two axes of the CMC joint.)



(b) **Top:** Custom data acquisition board. **Bottom:** Strain gauges attached to Castroviejo tool.

Fig. 2. Sensors and the board.

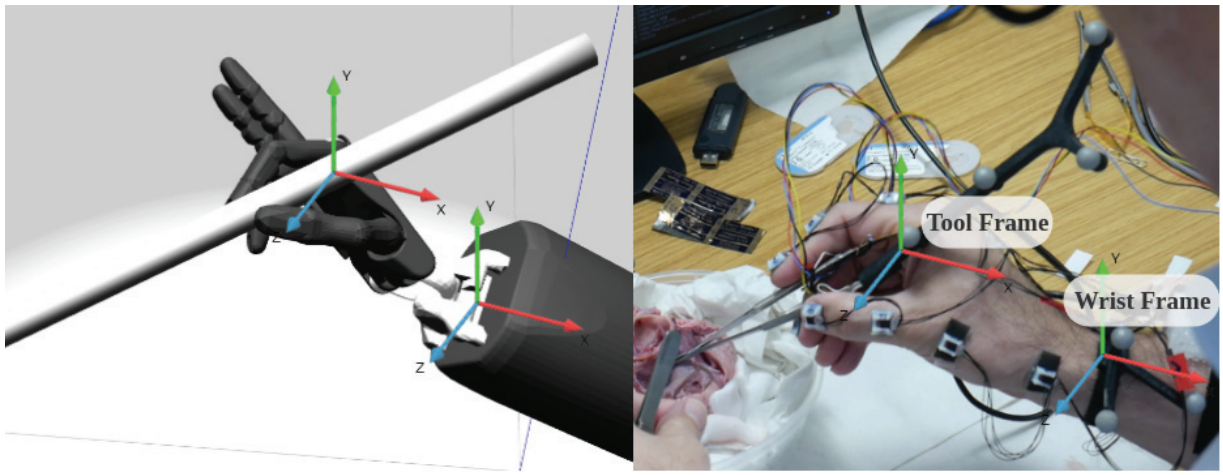


Fig. 3. **Right:** Hand pose and tool tracking using reflective markers, and **Left:** Gazebo simulation [23] of the acquired poses and orientations represented using the Shadow hand model [24].

Robotic Operating System (ROS) [25] on Ubuntu 16.04 for data recordings and processing.

A custom data acquisition and data transmission system was designed to capture finger and thumb joint angles. Since middle and index fingers and thumb have the key role in fine grasping and manipulation, we track their poses alongside the wrist using IMU sensors (BNO055 sensors from Bosch, see Fig. 2(a)). BNO055 chip has two different I2C bus address and we use PCA9548A 8-channel I2C switch to connect every pair of sensors to a single bus. Core of the system is a Micro-chip ATSAMD21 (Arm Cortex-M0+ processor) running on 48 MHz, which is responsible for setting a correct configuration of the I2C switch, reading the orientation of each sensor, and sending the values to the computer through a USB port. All data captured by the board are framed as JSON data structures. The micro-controller also initializes all the sensors and puts them into calibration mode, if required. I2C bus is operated at 400 KHZ frequency. IMU sensors produce quaternions and their relative values are calculated to obtain joint angles for each finger/thumb/wrist joint. These joint angles are then stored and visualized in Gazebo [23] and Rviz [26] before used as inputs into the mapping model.

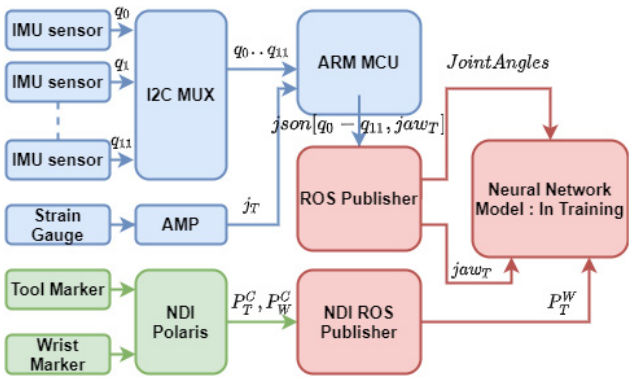


Fig. 4. System block diagram showing inputs of the neural network (joint angles) and outputs of the network (orientation of the surgical instrument and jaw angle of the instrument).

One of the limitations of tracking accuracy is the IMU drift. To solve this problem, we only used the relative angle between two IMUs therefore reducing the effect of drift. Another solution was to put the device into the calibration mode after each sequence of recording (approximately every 5 min).

Assuming $q_0 - q_{11}$ are corresponding quaternion values from the 12 IMU sensors, a relative quaternion for each joint can be calculated. For instance, q_0^1 represents q_0 with respect to q_1 can be found as follows:

$$q_0^1 = q_1 \otimes q_0^{-1}, \tag{1}$$

where q_0^{-1} is inverse quaternion and can be found by negating the w-component of quaternion:

$$q^{-1} = [q_x, q_y, q_z, -q_w]. \tag{2}$$

Calculating the relative angles in quaternion will help us to avoid problems associated with Euler angles. Final joint angles however are transformed to Euler angles because of representation and matching with other signals.

$$\begin{aligned}\text{Joint Angles} = [&ff_{j1}, ff_{j2}, ff_{j3}, ff_{j4}, mf_{j1}, mf_{j2},\\ &mf_{j3}, mf_{j4}, \ldots, th_{j1}, th_{j2}, th_{j3},\\ &th_{j4}, th_{j5}, wr_{j1}, wr_{j2}].\end{aligned} \tag{3}$$

Data from the four strain gauges are sampled using an ADC (HX711, AVIA Semiconductor [27]) with an on-chip active low noise programmable gain amplifier (PGA) with a selectable gain of 32, 64, and 128. $\text{jaw}_T(t)$ is jaw angle of castro tool which is represented in Euler. Assuming castro tool's jaw angle can vary between 30° and 0° in fully open and close scenarios, $\text{jaw}_T(t)$ can be calculated by linear mapping measurements from strain gauges.

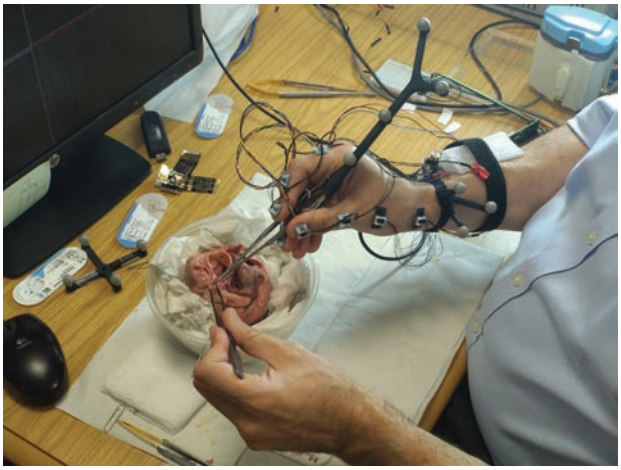
The main MCU synchronizes data from the strain gauges and the IMU and sends them to the PC at the rate of 50 Hz through the USB.

2.2. *Data collection*

The motion tracking data were collected from an experienced cardiac surgeon performing typical cardiac tasks (coronary and mitral valve repair surgery) on an *ex-vivo* animal heart, which included all the different instances of cutting, suturing and knotting. We made six recording sequences of different duration which all added up to 1 h.

The obtained datasets were pre-processed to filter out incomplete or missing data, e.g. occlusions of the tool or wrist markers that occasionally caused data loss. In the next stage, input and output values are normalized to get better performance from neural network. All 10,000 samples at the sample rate of 30 Hz in the dataset are timestamped values of hand joint angles, tool orientation, and tool jaw angle. The mock surgical experimental setup for the data collection is shown in Fig. 5 with a representative graph of the collected data shown in Fig. 6.

(a) Experimental setup with the optical tracker and camera.



(b) Hand tracking with the IMU chain and Polaris NDI marker tool during typical open heart surgery tasks on an ex-vivo porcine heart.

Fig. 5. Data collection.

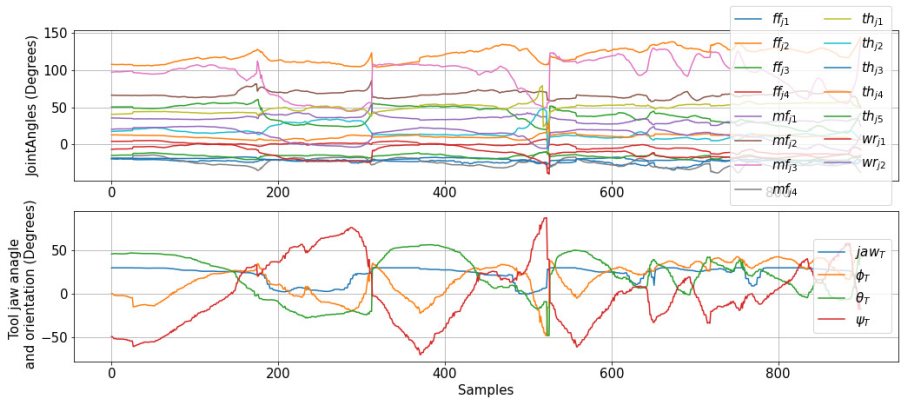


Fig. 6. Sample data collected during the surgery showing joints angles as input and tool orientation and tool jaw angle as output.

2.3. *Hand-tool modeling*

Due to the nonlinear multi-input-multi-output nature of the collected data, simple machine learning methods such as multivariate linear regression, or Support Vector Regression (SVR) do not provide satisfactory results. A Deep Neural Network (DNN), however, can be efficiently trained to map multi-DOF tool motion from the corresponding hand motion. Here we use two types of neural networks, Deep Feed-forward Neural Network (DFNN) and a Long Short-Term Memory (LSTM) neural networks, and compare their performance in terms of time and accuracy. Deep Feed-forward Neural Network (DFNN) is an FNN network that has more than two layers, which enables the network to learn more complex patterns of data. A Long-Short Term Memory (LSTM) neural network is an artificial Recurrent Neural Network (RNN) that has feedback connections between the layers which enables learning a sequence of data without facing limitations of the RNNs such as the vanishing gradient [28,29].

A DNN using Keras library in Python with n number of neurons in each hidden layer and l number of hidden layers is used here. We use both DFNN and LSTM neural networks and compare their performance in terms of time and accuracy.

Inputs of the neural network at a time of t are comprised of 15 hand joint angle inputs in Euler representation:

$$X(t) = [ff_{j1}, ff_{j2}, ff_{j3}, ff_{j4}, mf_{j1}, mf_{j2}, mf_{j3}, mf_{j4}, \ldots, th_{j1}, th_{j2}, th_{j3}, th_{j4}, th_{j5}, wr_{j1}, wr_{j2}], \quad (4)$$

(see Fig. 2(a) for notation) and output of the network during the training is comprised of

$$Y(t) = [P_T(t), \text{jaw}_T(t)] = [\phi_T(t), \theta_T(t), \psi_T(t), \text{jaw}_T(t)], \quad (5)$$

where P_T is surgical tool pose and jaw_t is surgical jaw angle.

After the neural network is adequately trained, the tracking markers and strain gauges attached on the tool are not required. The surgeon controls the tool by wearing the sensor chain while the neural network determines the outputs only from JointAngles(t). The outputs of the network are $[\phi_d(t), \theta_d(t), \psi_d(t), j_d(t)]$. This is one of the benefits of the proposed method as using vision-based tracking methods like Polaris would not work when the markers are covered, and would need a more complicated setup

Details of selected network configuration are as follows:

- Optimization method: Adam with default parameters (learning rate (lr)$= 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, decay $= 0.0$).
- Input neurons: 15 features (joint angles).
- Output neurons: 4 (3 neuron for tool orientation + 1 for jaw angle).
- Dataset size: 10,000 samples.
- Train test split: 80% for training, 20% for testing.
- Number of epoches: 200.
- Loss function: mse.
- FNN activation function: relu.

The rationale of these parameters was to achieve the highest precision in estimating the tool orientation with a smallest and therefore fastest network size.

We performed all the training and testing on a MacBook pro Core i7 with the following specs:

- 2.6 GHz 6-Core Intel Core i7.
- Radeon Pro 555X 4 GB Intel UHD Graphics 630 1536 MB.
- 16 GB 2400 MHz DDR4.

Loss function:

Loss function quantifies the error between output of the algorithm and the given target value. Mean Squared Error (MSE) is the loss function used here.

$$\text{Loss Function} = \frac{\sum_{i=1}^{n} (Y(i) - \hat{Y(i)})^2}{n}, \quad (6)$$

where $Y(i)$ is the given output and $Y(\hat{i})$ is the predicted output.

2.4. *Feature importance*

Feature importance identifies which joints are main contributors in creating certain types of tool motion. Decision Tree Regressor Gini importance or Mean Decrease in Impurity (MDI) [29] was used to establish a correlation between the hand joints and the tool variables.

2.5. *Principal component analysis*

We also used Principal Component Analysis (PCA) to extract five principal components of the dataset to train the two networks. LSTM networks are more complex than their DFNN counterparts due to the recurrent layers which consume more processing resources but they usually have better performance when it comes to time-series. The performed dimensionality reduction lowers the network complexity while maintaining the output accuracy, especially in LSTM networks. This is essential in teleoperation where latency plays a crucial role in the stability of teleoperation systems, specifically bi-directional teleoperation systems that require fast loop update rates.

3. Results and Discussion

In Fig. 7, we report mean squared error of estimated tool pose compared to the ground truth obtained using Polaris sensor for different network depth and layer size for both DFNN and LSTM architectures. As it can be seen, having more than two deep layers increases the network complexity and affects the processing time. However, it appears that this does not reduce the error further. The validity of the identified error for surgical robotics we compared to [30] states the motion trajectories in cardiac surgery for relevant movements to be 0.22 to 0.81 mm in the lateral plane (x/y-axis) and 0.52.6 mm out of the plane (z-axis). Our errors currently exceed the stated values. Thus, increasing the number of the neurons over 20 in each layer does not improve networks efficiency. Data shown in Fig. 8 demonstrate that the network can accurately estimate tool orientation and jaw angle while the results in Table 1 show an advantage of the LSTM neural network in terms of accuracy accompanied with

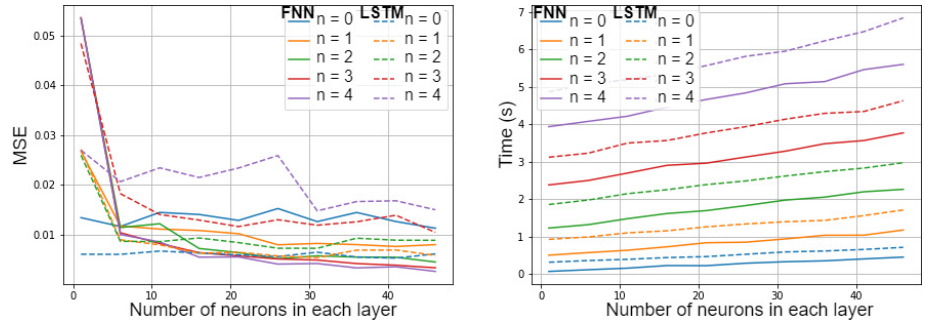


Fig. 7. Effect of the number of neurons in each layer (n) and the number of hidden layers (l) on estimation accuracy (mean squared error with normalized units) and execution time in seconds.

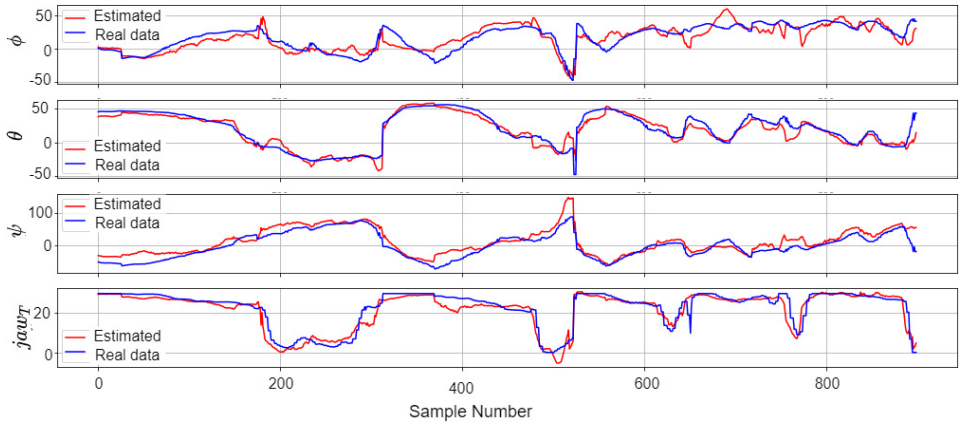


Fig. 8. (Color online) Surgical tool pose and jaw angle. Neural Network estimated values from unused data (red) and real data from optical sensor as ground truth (blue) $[\phi_T(t), \theta_T(t), \psi_T(t), \text{jaw}_T(t)]$ (the same slice as in Fig 6).

Table 1. Final results from LSTM and DFNN models.

Model	DFNN				LSTM			
joints	ϕ_T	θ_T	ψ_T	jaw_T	ϕ_T	θ_T	ψ_T	jaw_T
Root mean squared error	7.5°	7.0°	12.1°	3.1°	7.2°	6.1°	10.9°	1.8°
r2 score on training set	91%				92%			
r2 score on testing set	90%				91%			
Execution time	0.16 s				0.46 s			

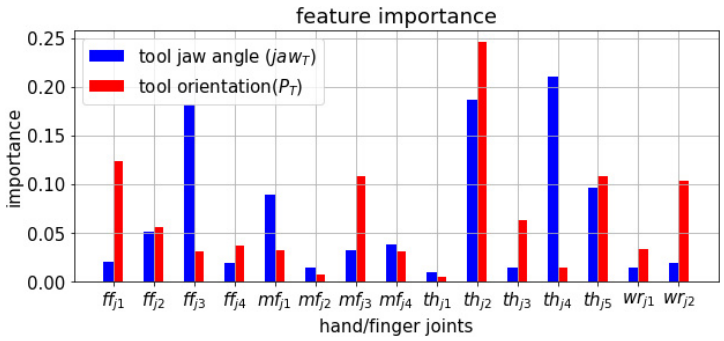


Fig. 9. Input feature importance with respect to corresponding outputs (tool orientation or jaw angle).

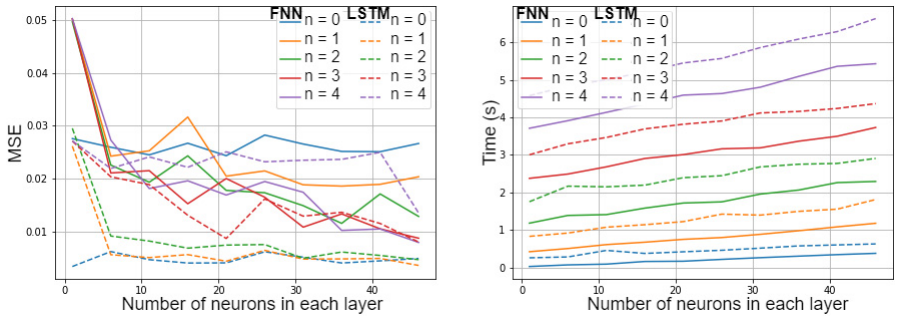


Fig. 10. Result of the neural network trained with reduced inputs (only five most important joints involved in surgery) (mean squared error with normalized units).

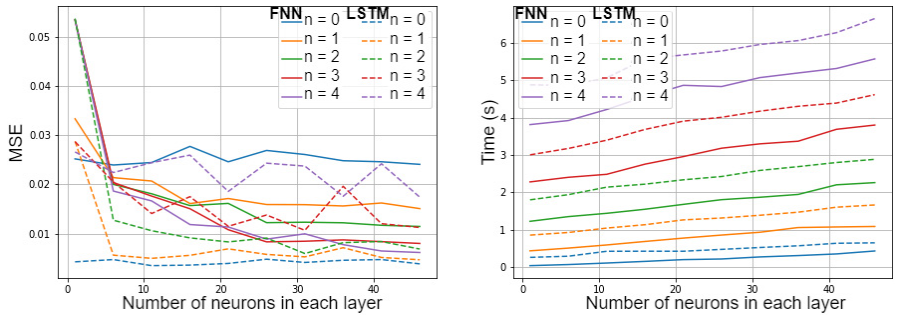


Fig. 11. Result of the neural network trained with PCA (five principal components as an input) (mean squared error with normalized units).

almost three times longer processing time. The close match between the predicted and real data values presented in Fig. 8 is encouraging considering that the predicted data were derived from the first prototype of our system, which was clearly meant to provide proof of concept data. We are confident that the existing discrepancies can be reduced markedly with the refinement of our prototype.

The importance of the input features for each output variable is shown in Fig. 9. It can be seen that th_{j2}, th_{j4}, and ff_{j3} are most important features for jaw opening of the tool while the tool orientation largely dependence on the th_{j2}, wr_{j2}, and ff_{j1}. This analysis also highlighted the less important joints such as mf_{j2} which impact on teleoperation of the surgical instrument can be mostly neglected.

When the two networks were trained and tested using only five most important joints, LSTM network demonstrated a better performance compared to DFNN as shown in Fig. 10. These findings can be particularly useful in designing a new system with a reduced number of tracked joints and interfacing hardware complexity.

Figure 11 shows the neural network trained with five principal components of the dataset. The LSTM neural network demonstrated slightly better performance in estimating the tool orientation.

4. Conclusion and Future Work

In this paper, we tested the hypothesis that motion of surgical instruments that perform fine movements in cardio surgery can be mapped to the hand/wrist motion of the surgeon. However, more work is needed to refine this approach and confirm its potential to enhance teleoperation. A wearable nonobstructive hand/wrist and surgical tool tracking system that can collect accurate multi-point data were built and tested. We created a mock up cardiac surgery test-bed where an experienced cardio surgeon performed typical surgical tasks. The collected datasets of hand/wrist joints as inputs and tool motion as outputs were used to train two types of DNNs — LSTM and DFNN. We compared performance of the two networks using all captured inputs but also identified importance of each hand joint on the tool motion. Performance of the two networks was again compared using the smaller number of salient inputs to reduce the network complexity. The implemented optimization allows the use of fewer joint angles as control inputs in order to achieve the same output performance. This means that the surgeon could teleoperate a Castroviejo-like instrument using the same type of hand movements as in open access surgery, potentially decreasing/eliminating surgeon's cognitive and muscular fatigue currently experienced with teleoperated surgical robots. Developing this approach further can address the need for surgical retraining to undertake same procedures through a completely different set of movements required in surgical robot teleoperation. Teleoperating a surgical robot using a wearable hand tracking system, combined with a VR headset would, as presented in [31], provides the surgeon with a unique ability to operate the robot closer to the patient and operating theatre team or from any remote sit [4].

Future advances in implementing the wearable tracking to a wider range of fine motion surgical instruments for cardiac and vascular surgeries would help establish and test the wearable tele-operation control concept. We have begun extending this framework to other surgical areas like arthroscopy with the similarly sized operating fields. While these results are encouraging, more relevant data are necessary to implement machine learning techniques which focuses us on conducting further user studies in this and other surgical scenarios. Another important area to focus would be smoothness of the predicted signal and effect of the signal smoothness and noise on accurate performance in microsurgical robotics.

Acknowledgments

This work was performed in SMARTsurg project supported by the European Unions Horizon 2020 research and innovation programme under Grant No. 732515.

References

1. H. Alemzadeh, J. Raman, N. Leveson, Z. Kalbarczyk and R. K. Iyer, Adverse events in robotic surgery: A retrospective study of 14 years of fda data, *PloS One* **11**(4) (2016) e0151470.
2. E. Rodriguez and W. Chitwood, Robotics in cardiac surgery, *Scand. J. Surg.* **98**(2) (2009) 120–124.
3. E. Quint and G. Sivakumar, The role of robotic technology in cardiac surgery, *Univ. Western Ontario Med. J.* **87**(2) (2018) 40–42.
4. A. Simorov, R. S. Otte, C. M. Kopietz and D. Oleynikov, Review of surgical robotics user interface: what is the best way to control robotic surgery?, *Surg. endosc.* **26**(8) (2012) 2117–2125.
5. L. S. Mattos, N. Deshpande, G. Barresi, L. Guastini and G. Peretti, A novel computerized surgeon–machine interface for robot-assisted laser phonomicrosurgery, *Laryngoscope* **124**(8) (2014) 1887–1894.
6. A. Mewes, B. Hensen, F. Wacker and C. Hansen, Touchless interaction with software in interventional radiology and surgery: A systematic literature review, *Int. J. Comput. Assist. Radiol. Surg.* **12**(2) (2017) 291–305.
7. G. R. Sutherland, S. Wolfsberger, S. Lama and K. Zarei-nia, The evolution of neuroarm, *Neurosurgery* **72**(1) (2013) A27–A32.
8. M. Fattahi Sani, S. Abeywardena, E. Psomopoulou, R. Ascione and S. Dogramadzi, Towards finger motion tracking and analyses for cardiac surgery, in *2019 15th Mediterranean Conf. Medical and Biological Engineering and Computing (MEDICON)* (Springer, 2019), pp. 1–11.
9. Y.-C. Du, C.-B. Shih, S.-C. Fan, H.-T. Lin and P.-J. Chen, An IMU-compensated skeletal tracking system using kinect for the upper limb, *Microsyst. Technol.* **24**(10) (2018) 4317–4327.
10. R. Li, Z. Liu and J. Tan, A survey on 3d hand pose estimation: Cameras, methods, and datasets, *Pattern Recognit.* **93** (2019) 251–272.
11. D. Lu, K. Xu and D. Huang, A data driven in-air-handwriting biometric authentication system, in *2017 IEEE Int. Joint Conf. Biometrics (IJCB)* (IEEE, 2017), pp. 531–537.
12. C. E. Reiley, E. Plaku and G. D. Hager, Motion generation of robotic surgical tasks: Learning from expert demonstrations, in *2010 Annual Int. Conf. IEEE Engineering in Medicine and Biology* (IEEE, 2010), pp. 967–970.

13. M. Power, H. Rafii-Tari, C. Bergeles, V. Vitiello and G.-Z. Yang, A cooperative control framework for haptic guidance of bimanual surgical tasks based on learning from demonstration, in *2015 IEEE Int. Conf. Robotics and Automation (ICRA)* (IEEE, 2015), pp. 5330–5337.
14. M. J. Fard, S. Ameri, R. B. Chinnam, A. K. Pandya, M. D. Klein and R. D. Ellis, Machine learning approach for skill evaluation in robotic-assisted surgery, preprint (2016), arXiv:1611.05136.
15. S. Speidel, M. Delles, C. Gutt and R. Dillmann, Tracking of instruments in minimally invasive surgery for surgical skill analysis, in *Int. Workshop Medical Imaging and VirtualReality* (Springer, 2006), pp. 148–155.
16. A. Ometov et al.et al., A survey on wearable technology: History, state-of-the-art and current challenges, *Comput. Netw.* **193** (2021) 108074.
17. B. D. Itkowitz, S. P. DiMaio and K. Y. Bark, Master finger tracking device and method of use in a minimally invasive surgical system (2013), US Patent 8,543,240.
18. B. D. Itkowitz, S. P. DiMaio and T. Zhao, Method and apparatus for hand gesture control in a minimally invasive surgical system (2015), US Patent 8,996,173.
19. S. Abeywardena, E. Psomopoulou, Efi amd Fattahi Sani, A. Tzemanaki and S. Dogramadzi, Control of a da vinci endowrist surgical instrument using a novel master controller, in *2019 15th Mediterranean Conf. Medical and Biological Engineering and Computing (MEDICON)* (Springer, 2019), pp. 1545–1550.
20. W.-Y. Loh, Classification and regression tree methods, *Wiley StatsRef: Statistics Reference Online* (2014).
21. N. Polaris, Northern digital polaris tracking system (2004).
22. A. Deguet, R. Kumar, R. Taylor and P. Kazanzides, The cisst libraries for computer assisted intervention systems, in *MICCAI Workshop Systems and Architectures for Computer Assisted Interventions*, Vol. 71 (2008).
23. N. Koenig and A. Howard, Design and use paradigms for gazebo, an open-source multi-robot simulator, in *2004 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)* (IEEE Cat. No. 04CH37566), Vol. 3 (IEEE, 2004), pp. 2149–2154.
24. Shadow Dexterous Hand (2020), www.shadowrobot.com/products/dexterous-hand, (-).
25. Stanford Artificial Intelligence Laboratory, Robotic operating system.
26. D. Hershberger, D. Gossow and J. Faust, Rviz, 3d visualization tool for ros, Vol. 1 (2019), http://wiki. ros. org/rviz.
27. D. HX711, 24-bit analog-to-digital converter (ADC) for weigh scales, AVIA Semiconductor (2014).
28. R. E. da Silva, J. Ondrej and A. Smolic, Using LSTM for automatic classification of human motion capture data, in *14th Int. Conf. Computer Vision, Imaging and Computer Graphics Theory and Applications*, Prague, Czech Republic, 2019, pp. 236–243.
29. L. Breiman, Random forests, *Mach. Learn.* **45**(1) (2001) 5–32.
30. P. Cattin, H. Dave, J. Grünenfelder, G. Szekely, M. Turina and G. Zünd, Trajectory of coronary motion and its significance in robotic motion cancellation, *Eur. J. Cardio-thoracic Surg.* **25**(5) (2004) 786–790.
31. A. Danioni, Study on dexterity of surgical robotic tools in a highly immersive concept (2020).