Original Study

ANNALS OF SURGERY OPEN

OPEN

Force Profile as Surgeon-Specific Signature

Amir Baghdadi, PhD,* Eddie Guo,* Sanju Lama, MD, PhD,* Rahul Singh, BSc,*
Michael Chow, MD, FRCSC,† and Garnette R. Sutherland, MD, FRCS, DSc. (Hon)*

Objective: To investigate the notion that a surgeon's force profile can be the signature of their identity and performance.
Summary background data: Surgeon performance in the operating room is an understudied topic. The advent of deep learning methods paired with a sensorized surgical device presents an opportunity to incorporate quantitative insight into surgical performance and processes. Using a device called the SmartForceps System and through automated analytics, we have previously reported surgeon force profile, surgical skill, and task classification. However, an investigation of whether an individual surgeon can be identified by surgical technique has yet to be studied.
Methods: In this study, we investigate multiple neural network architectures to identify the surgeon associated with their time-series tool-tissue forces using bipolar forceps data. The surgeon associated with each 10-second window of force data was labeled, and the data were randomly split into 80% for model training and validation (10% validation) and 20% for testing. Data imbalance was mitigated through subsampling from more populated classes with a random size adjustment based on 0.1% of sample counts in the respective class. An exploratory analysis of force segments was performed to investigate underlying patterns differentiating individual surgical techniques.
Results: In a dataset of 2819 ten-second time segments from 89 neurosurgical cases, the best-performing model achieved a micro-average area under the curve of 0.97, a testing F1-score of 0.82, a sensitivity of 82%, and a precision of 82%. This model was a time-series ResNet model to extract features from the time-series data followed by a linearized output into the XGBoost algorithm. Furthermore, we found that convolutional neural networks outperformed long short-term memory networks in performance and speed. Using a weighted average approach, an ensemble model was able to identify an expert surgeon with 83.8% accuracy using a validation dataset.
Conclusions: Our results demonstrate that each surgeon has a unique force profile amenable to identification using deep learning methods. We anticipate our models will enable a quantitative framework to provide bespoke feedback to surgeons and to track their skill progression longitudinally. Furthermore, the ability to recognize individual surgeons introduces the mechanism of correlating outcome to surgeon performance.

Keywords: cloud computing, deep learning, surgeon signature skill, time-series modeling, tool-tissue interaction force

INTRODUCTION

In reporting surgical outcomes, publications often emphasize patient-related variables, including disease severity, co-morbidities, and physiological parameters.[1–4] Less studied is the correlation between surgeons, objective analysis of surgical skills, and their relationship with patient outcome.[5–7] In surgical skill acquisition, trainees have often recognized master surgeons who have unique technical skills that correlate with intra- and perioperative success. Although the inherent attributes and experience of a surgeon may account for these differences, quantitative analysis of such remains elusive. With the advent of artificial intelligence and data modeling on sensor-enabled surgery tools, the following questions arise: can a data-driven force profile quantify a surgeon's technique or finesse, adding an objective parameter to predict intra- and perioperative variability? And can a surgeon's force profile be the surgeon's signature and identity?

In robot-assisted surgery, ensuring surgeon identification is important for secure signal transmission as a requisite for safe and successful surgery. One safeguard can be anomaly detection through machine learning of a "surgical signature" for the primary surgeon to impede the alteration of robotic manipulation by a malevolent hacker. Deviation of robot maneuver from the movement patterns of the primary surgeon can trigger an identity verification process before authorizing further actions.[8,9] Static and dynamic physical traits (eg, body shape, eye movement, and gait features) are recognized as biometrical signatures for person detection that can be measured through visual and motion sensors.[10–14] Similarly, tool orientation and dynamics, such as velocity and tissue force, for each surgeon can construct a signature correlating with the surgeon's skill level.[9,15]

In our previous studies, we developed a novel data framework and application, called the SmartForceps System, which creates a digitized environment for real-time intraoperative monitoring, recording, and secure upload of surgical procedures.[16–19] Our operating room data intelligence paradigm uses the characteristics of

From the *Project neuroArm, Department of Clinical Neurosciences, and Hotchkiss Brain Institute University of Calgary, Calgary, Alberta, Canada; and †Department of Surgery, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Canada

This study has received the support of the Alberta-Germany Collaborative Funds (German Canadian Centre for Innovation and Research) and the Canadian Institutes of Health Research (Grant No. 390405)-Technology Commercialization.

All authors are associated with a university spin-off company called OrbSurgical Ltd., based in Calgary, AB. In addition to intellectual property (IP) arising from Project neuroArm-University of Calgary, the company also hosts SmartForceps System IP. The other authors declare that there is nothing to disclose

A GitHub repository contains de-identified data and model code: https://github.com/smartforceps/ai_models.

Reprints: Garnette R. Sutherland, MD, FRCS, DSc. (Hon), Department of Clinical Neurosciences, and Hotchkiss Brain Institute University of Calgary, Calgary, Alberta, Canada. E-mail: garnette@ucalgary.ca.

Annals of Surgery Open (2023) 3:e326

Received: 27 January 2023; Accepted 22 July 2023

Published online 15 September 2023

DOI: 10.1097/AS9.0000000000000326

tool-tissue interaction force and surgical maneuvers as an objective metric for assessing and reporting surgical competency compared with an expert surgeon. Accurate data collection of individual surgeons and surgeon identification stratified by skill provides a richer context for an individualized performance report. This study builds upon the SmartForceps System by investigating the notion that a surgeon's force profile can be the signature of their identity and performance.[17–21] Here we present evidence for a surgeon identification system to catalog a correlational linkage between surgical finesse and surgeon through 3 classes of deep learning algorithms.

METHODS

This study was conducted to determine whether machine-learning algorithms could predict individual surgeons using time-series surgical tool-tissue force data. Predictions were made using 3 classes of machine-learning algorithms: (1) recurrent neural network, (2) convolutional neural network (CNN), and (3) CNN as a feature selector for a gradient-boosted tree. The models were validated by applying the top-performing model to a prospective dataset spanning 2 years across 2 surgical centers.

Data Recording and Cloud Analysis

The SmartForceps System (developed at Project neuroArm, University of Calgary, Calgary AB, Canada) displays and records real-time tool-tissue force data. Audio recordings of each surgeon's voice accompanied force recordings, which indicated the duration of force application and specific task names. Force recordings included tool-tissue forces from the left and right prong of a sensorized bipolar forceps called the SmartForceps at a 20 Hz sampling rate. A supervised dataset was created from this data by labeling each force recording. Health Canada (ITA 329641 Class II, 2021) approved the technology used in this study, which was reviewed and approved by the Conjoint Health Research and Ethics Board of the University of Calgary, Calgary, Alberta, Canada (REB19-0114). A detailed description of the technology development, the preclinical trial, and the clinical phase of the study has been published.[17–23]

An analysis of tool-tissue interaction force data was conducted on 89 neurosurgery cases (50 cases with manual labeling of data, 39 patients with automated predicted labels based on previously developed machine-learning models), consisting of 18 meningioma, 20 glioma, 4 hemangioblastoma, 22 schwannoma, 1 carotid plaque, 1 choroid plexus papilloma resection, and 23 miscellaneous cases. These recordings corresponded to 7.8 hours of tool-tissue forces, with 563,849 force data points. Six surgeons performed the cases: 1 surgeon with more than 30 years of experience and 3 final year neurosurgical residents with post-graduate years level exceeding 5 years from Foothills Hospital, Calgary; 1 surgeon with 19 years of experience and 1 final year neurosurgical resident from the University of Alberta Hospital, Edmonton.

Cloud architecture compliant with the Health Insurance Portability and Accountability Act and Personal Information Protection and Electronic Documents Act regulations were implemented to retain and process intraoperative de-identified data in transit using transport layer security and at rest using advanced encryption standard (AES-256) through reliance on Microsoft Azure platform (Microsoft USA). This further authorized the use of organizational credentials for secure authentication. We also developed web and mobile applications for monitoring and analyzing force-related data/features, which can be accessed at smartforceps-app.azurewebsites.net.

Workflow Architecture

A workflow architecture was developed and implemented to create data analytics and surgeon signature identification models (Fig. 1). It provided a framework for developing surgeon recognition models through cloud implementation for data warehousing and preprocessing, including deep learning-based segmentation and feature engineering. Our framework can model nonstationary time-series data with changing mean, variance, and frequency characteristics without assuming underlying patterns in force data.

Surgeon Signature Recognition Model

Using deployed models of phase 1 analytics of the SmartForceps platform,[18] force profiles of individual surgeons were identified using a combination of manually and automatically generated segments with surgeon ID labels. Post-hoc tags were applied to surgeon data to correct data records from multiple surgeons using the device in 1 surgical case. All time-series data were standardized and split into 200-point windows by the surgeon, corresponding to 10 seconds of recorded forces. This created 472, 468, 476, 466, 471, and 466 force segment windows for surgeons 1 to 6, respectively. These surgeons associated with these windows were encoded as one-hot vectors, and the data were randomly split into 80% for model training and validation (10% validation) and 20% for testing. Data imbalance, as a challenge in creating prediction bias toward majority classes, was mitigated through subsampling from more populated classes with a random size adjustment based on 0.1% of sample counts in the respective class. Exploratory analysis of force segments was performed to investigate underlying patterns differentiating individual surgical techniques.

To recognize and classify each surgeon, we analyzed multiple deep neural networks; we created a long short-term memory (LSTM) network, CNN, ResNet,[24] InceptionTime,[25] and Force Time-series Feature-based InceptionTime (FTFIT) network. The LSTM network consisted of a single LSTM layer with 100 LSTM units followed by 2 dense layers for classification. The CNN consisted of a convolutional layer followed by max pooling, another convolutional layer followed by max pooling, and then a dense layer for classification. All models were trained using backpropagation with the Adam optimizer. The ResNet architecture was inspired by previously published work.[26] Following this analysis, the convolutional components of ResNet were used as a feature selector for the XGBoost algorithm to identify individual surgeons; we define this model as ResNet-XGBoost (Fig. 2). To improve the prediction power of the FTFIT model, a subset of 29 hand-crafted features (eg, descriptive statistics, heterogeneity, and entropy) were calculated for each window of 200 data points and added as a third feature. These time-series features were extracted from segmented data after noise reduction, and the data were resampled to match a 200-point force data window, corresponding to 10 seconds of recorded forces.

For each machine-learning model, the following metrics were obtained: loss and accuracy for both training and validation data in each epoch; classification report including accuracy, sensitivity/recall, precision, F1-score, and area under the curve (AUC) for receiver operating characteristic. Models were trained on a Lambda Graphics Processing Unit workstation with an Intel Core i9-9820X (10 cores, 4.20 GHz turbo) Central Processing Unit, 2 Titan RTX with NVLink Graphics Processing Units, and 64 GB of memory.

Validation Experiment

The validation dataset for the expert neurosurgeon was recorded in May 2023 (ie, 2 years after the original data collection between October 2021 and May 2023). The best-performing model from the surgeon signature recognition task was used to create 5 binary prediction models that compare Surgeon

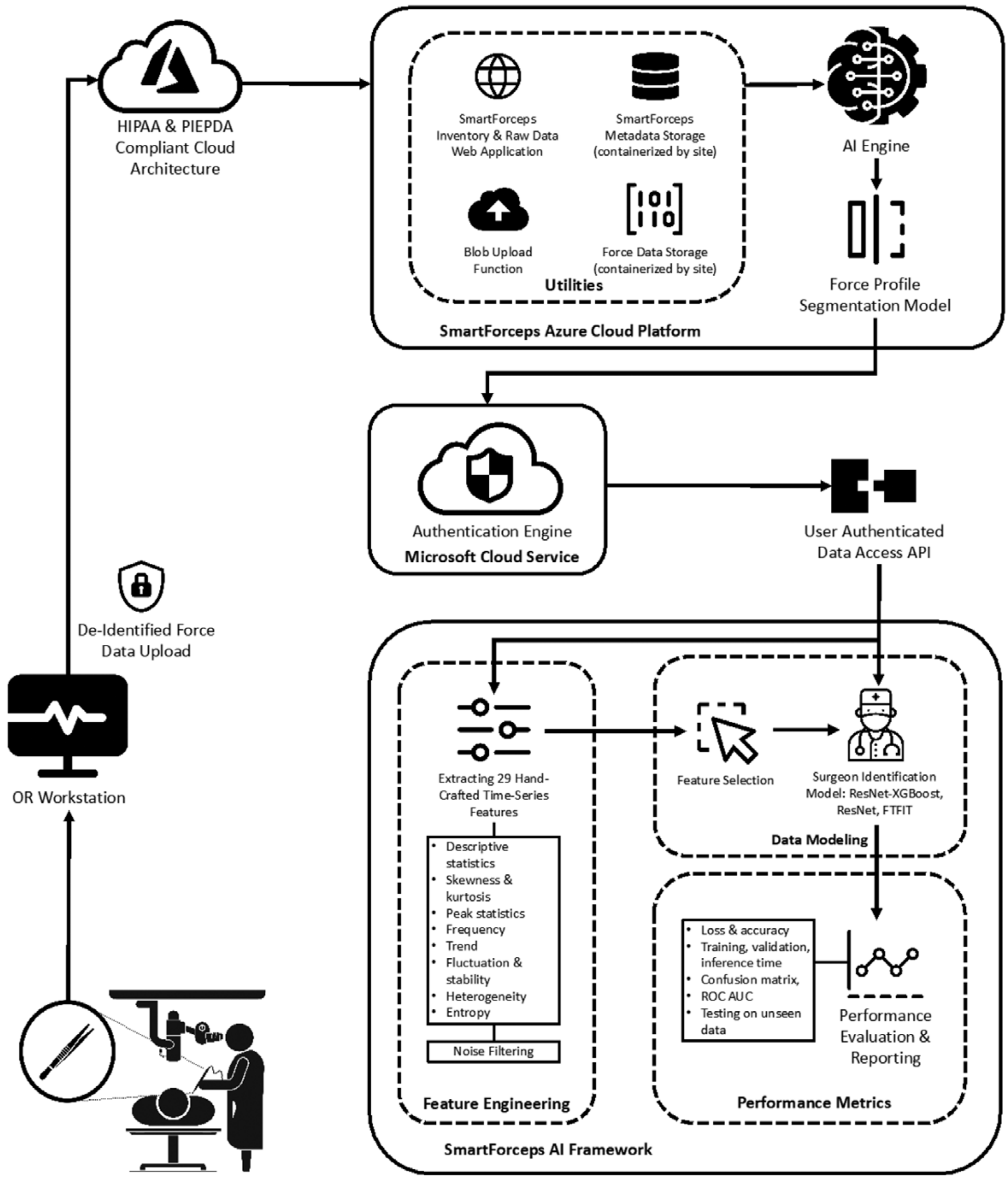

FIGURE 1. Workflow architecture of SmartForceps platform for surgeon identification. A HIPAA- and PIEPDA-compliant platform was used to store and analyze the forces of tool-tissue interaction. As part of the artificial intelligence (AI) modeling architecture, InceptionTime, ResNet, and ResNet-XGBoost models were utilized for surgeon identification, followed by performance evaluation reports. Visualization was created with icons obtained from https://www.iconfinder.com. HIPAA indicates Health Insurance Portability and Accountability Act; PIEPDA, Personal Information Protection and Electronic Documents Act.

1 against Surgeons 2 to 6. All 5 model outputs were combined into a single ensemble model by multiplying their respective model probabilities with experimentally determined weights. The weights w_{1i} were set *a priori* to the weighted average accuracies obtained in training. The predicted surgeon was obtained by using an argmax operation on the weighted probabilities:

$$\text{weighted} \quad \text{probability} = \sum_{i=2}^{6} w_{1i} \cdot (\text{model} \quad \text{probability})_{1i}$$

$$\text{predicted} \quad \text{surgeon} = \text{argmax}\,(\text{weighted} \quad \text{probability})$$

RESULTS

Force profile segments of each studied surgeon were visualized using force range, entropy, and maximum force. The scatter plot of data point distributions shows a distance-based density change across the force range and entropy (surgeon 1 = 1.98, surgeon 2 = 4.48, surgeon 3 = 1.29, surgeon 4 = 6.13, surgeon 5 = 3.03, and surgeon 6 = 2.02). A positive correlation (ρ = 0.91) can also be seen between force range and maximum force (Fig. 3).

The best performance was achieved using the ResNet-XGBoost model with a micro-average AUC of 0.97 (Fig. 2). Model parameters included a learning rate of 0.001 and patience of 10 for early stopping with 200 moving windows and batch sizes of 128 (Table 1). At epoch 86 of 200, the model reached

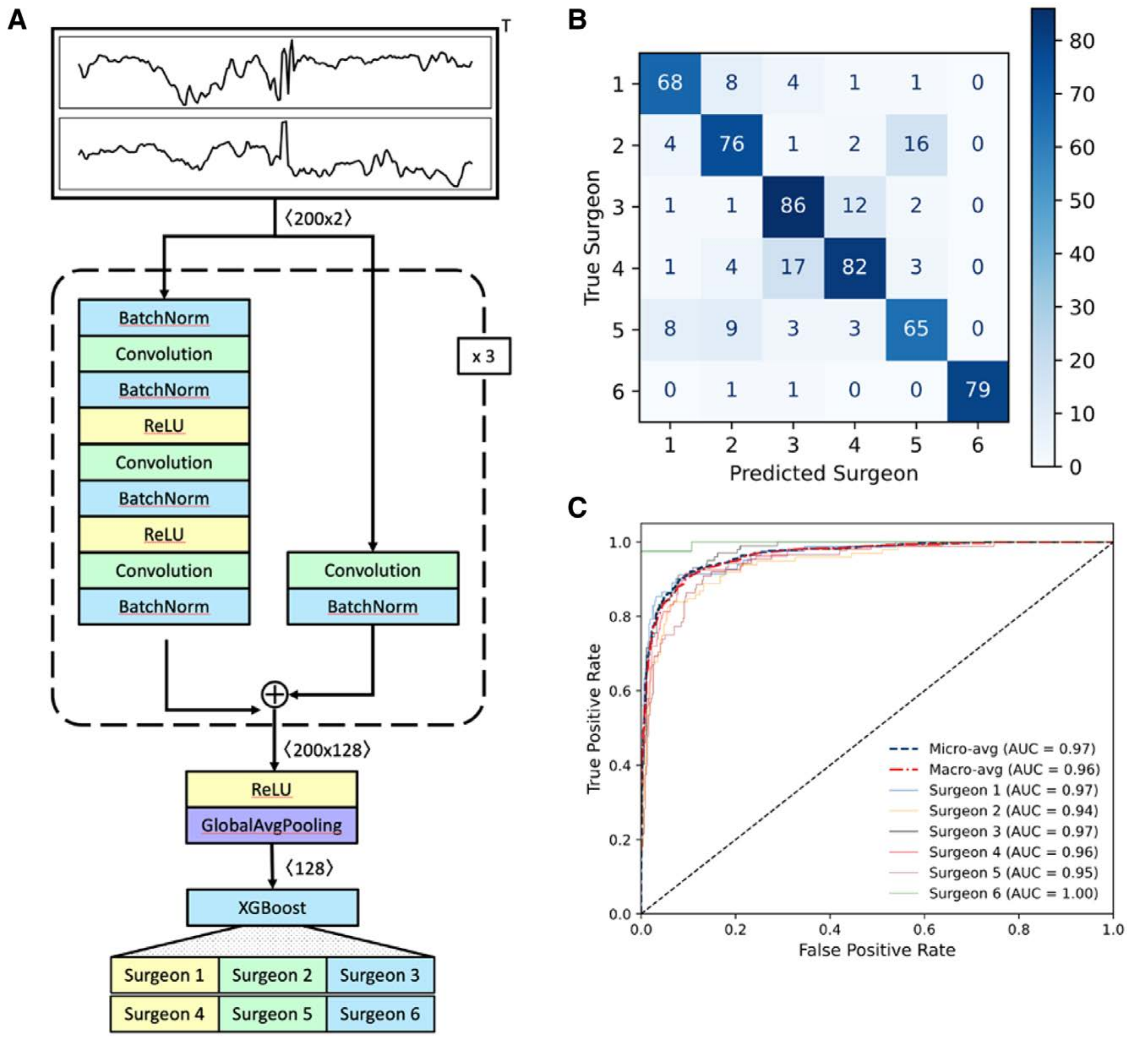

FIGURE 2. Performance of ResNet-XGBoost model for surgeon recognition. **A,** ResNet-XGBoost model architecture. Our deep neural network consisted of 12 convolutional layers followed by a linear output into the XGBoost algorithm, which predicts the surgeon associated with each segment. The network inputs are 200 × 2 force segments, corresponding to the left and right bipolar prong force data. B, ROC curves of the ResNet-XGBoost algorithm on the testing dataset. C, Confusion matrix metrics of the ResNet-XGBoost algorithm on the testing dataset. T = transpose; all convolutional layers use 1D filters with two channels. ROC indicates receiver operating characteristic.

minimum validation loss (validation loss = 0.5503 and training loss = 0.3123). The inference time on 564 test samples was 0.76 seconds. For unseen instances (ie, testing dataset) of force data, the macro-average AUC of receiver operating characteristic was 0.96, and the accuracy and weighted F-score both were 0.82 (Table 1).

In validation experiments, ResNet-XGBoost was used for the ensemble binary model. The ensemble weights, set a priori to the weighted average accuracies obtained in training, were 0.73, 0.73, 0.71, 0.58, and 0.97, for surgeons 1 *versus* 2 to 6, respectively. The results showed an accuracy of 83.8% in identifying surgeon 1 among others, visualized in Fig. 4.

DISCUSSION

This study provides proof-of-concept evidence for a surgeon identification system defined by bipolar forceps tool-tissue force profile. We provide evidence that deep learning models can identify individual surgeons given their time-series force data. In particular, the ResNet-XGBoost (AUC = 0.97), InceptionTime (AUC = 0.96), and ResNet (AUC = 0.95) models offer robust performance in identifying individual surgeons. These results suggest that the force profile characterizes a surgeon's signature and identity. They further support that profiles of tool-tissue force applications can be analyzed using a machine-learning approach. We thus present an original approach and finding for a surgeon identification system that catalogs correlational linkage between surgical finesse and surgeon, offering an automated platform to identify the nuances of surgery and surgeons by skill level.

A feature-based analysis of force profiles revealed a meaningful difference among surgeons for entropy, force range, and maximum force ($P < 0.05$). Tukey Honest Significant Difference tests showed significant differences among the binary combination of surgeons in the 3 features except for surgeon 1 *versus* 4 and surgeon 3 *versus* 6 in entropy, surgeon 2 *versus* 3, 5, and 6 and surgeon 6 *versus* 3 and 5 in maximum force, and surgeon 2 *versus* 5 and 6 and surgeon 5 *versus* 6 in force range. The positive correlation of maximum force and force range for surgeons emphasizes

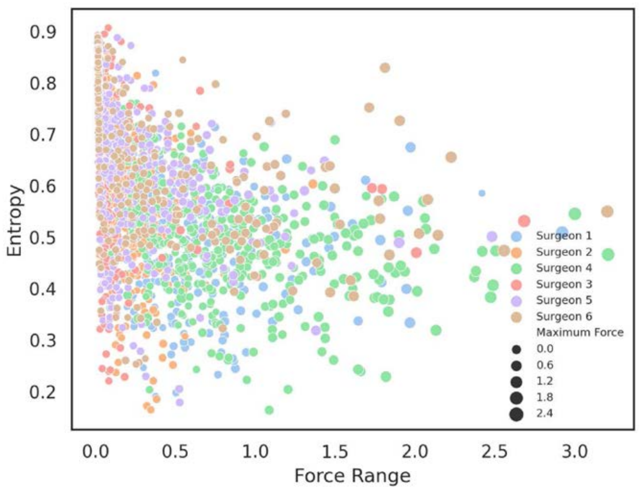

FIGURE 3. Distribution of surgeons' force segments visualized using force range, entropy, and maximum force. Data point densities for each surgeon change over force range and entropy, and a positive correlation exists between range force and maximum force.

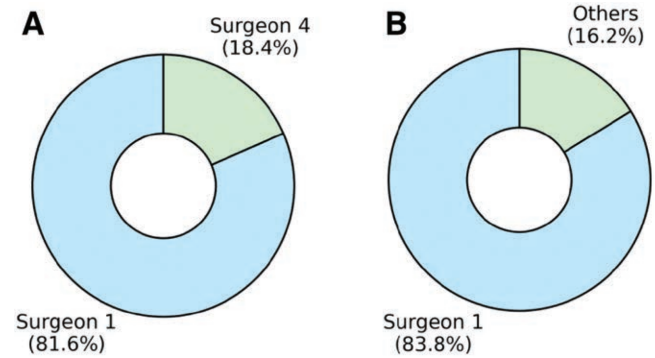

FIGURE 4. Distribution of surgeon data and predictions in the validation dataset used in ensemble-based binary ResNet-XGBoost model. Identifying surgeon 1 was >80% accurate in the validation dataset collected 2 years after the training phase. A, The validation dataset was predicted to be surgeon 1 in 81.6% of cases. B, In this figure, 83.8% of the validation dataset was predicted as surgeon 1.

that the higher force range is from positive force application (eg, coagulation) rather than negative forces (eg, retraction). The significance of feature comparison results shows a potential avenue for further exploration of fusing custom-designed features into deep learning models.

Utilizing time-series data, machine-learning models were developed by fusing custom-designed features with the receptors of the InceptionTime network, that is, the FTFIT deep learning model. However, for the present data, the highest performance belonged to models without feature infusion, that is, ResNet and ResNet-XGBoost. Our results show that ResNet-XGBoost performed with the highest accuracy of 82% (F1-score of 82%, weighted average of 82%, AUC micro average of 0.97, and macro average of 0.96). Furthermore, models ingesting standardized data trained significantly faster and had comparable or higher accuracy than models trained on unstandardized data.

The performance of the model was evaluated using binary classifiers between various surgeons' historical data and future data recordings. The slight decline in the accuracy from the ResNet-XGBoost model (acc = 82%) for all surgeons to our binarized ensemble model (acc = 84%) suggests that individual surgical technique transforms gradually over time. Based on our results, surgeon prediction models should be retrained every few months to account for these changes and optimize their performance over time.

To the best of our knowledge, this is the first article to automate the identification of individual surgeons. Previous methods focusing on person recognition using gesture data have been reported.[27,28] In contrast to our study, previous signature identification research primarily investigated security and safety to avoid unintended intrusions. For instance, 1 group developed RFnet, that is, a multi-branch 1D-CNN network, for classifying gestures using time-series data.[28] The system could identify an authorized person with a potential application in smart homes for access prevention or personalizing privileges.[28] Health and fitness data collected from wearable devices has been studied as a unique identifier of a person, which can pose a potential threat to targeted advertisement and violation of privacy rights.[27] Additionally, current literature on using deep learning methods in surgery sought to characterize surgeon skill level and surgical task identification using one or a combination of video object tracking or detection, spatiotemporal video descriptors, robotic kinematics, and virtual reality interfaces.[29–35] Furthermore, limitations of previous studies include the use of the da Vinci Surgical System on bench-top surgical training models rather than direct manipulation of the surgical tools during clinical cases.[36] In contrast to these studies, we use a unique data modality, surgical tool-tissue force time-series data, to automatically identify individual surgeons.

Identification of individual surgeons provides the opportunity to include this factor in explaining or accounting for variability in patient outcome.[5–7,37–39] In so doing, this study introduces a quantitative paradigm to identify the defining features of individual surgeons and their surgical performance. Using this framework, we could quantify some of the nuances of surgical technique for the first time. Such a methodology that links tool-tissue force profiles corresponding to the surgeon via machine learning can

TABLE 1.

Performance of Deep Learning Models for Surgeon Recognition

Metric	LSTM	CNN	FTFIT	ResNet	InceptionTime	ResNet-XGBoost
Best-performing hyperparameters	*lr = 0.001 Window = 200 batch size = 128	lr = 0.001 window = 200 filters/layer = 64 batch size = 128	lr = 0.0001 depth = 6 window = 200 batch size = 16	lr = 0.001 window = 200 batch size = 128	lr = 0.0001 depth = 6 window = 200 batch size = 16	lr = 0.001 window = 200 batch size = 128
Mean training time per epoch	13.7 s	0.23 s	2.30 s	2.54 s	2.00 s	1.40 s
Inference time (on 469 test samples)	10.0 s	0.14 s	0.25 s	1.53 s	0.49 s	0.76 s
Accuracy	0.55	0.63	0.69	0.74	0.78	0.82
Sensitivity/recall	0.55	0.63	0.68	0.75	0.78	0.82
Precision	0.59	0.64	0.68	0.74	0.78	0.82
Testing F1-score	0.54	0.63	0.69	0.74	0.78	0.82
Micro-average AUC	**0.86**	**0.88**	**0.93**	**0.95**	**0.96**	**0.97**

*lr = learning rate.

AUC indicates area under the curve; CNN, convolutional neural network; FTFIT, Force Time-series Feature-based InceptionTime; LSTM, long short-term memory.

well be the surgeon's signature. Such a signature, quantitative and traceable, may provide a true basis for linking patient outcome to the surgeon operating on a given procedure.

The strengths of this study include the incorporation of a novel data modality for a novel predictive task and the variety of state-of-the-art deep learning algorithms used. This included trainees in their final year, mentored by an attending surgeon at each site, thus creating a representative sample of force profiles across 2 sites. The models presented here were able to uniquely discern these individual surgeons. Limitations of this study include the small sample size of surgeons, given the early stage of technology integration. Although 89 cases were performed, more data is required to confirm the scalability of deep learning models in intelligently detecting the force profile of individual surgeons.

Ongoing work aims to validate our deep learning models using an expanded dataset with more surgeons and to describe the most predictive features which differentiate individual surgeons. Of interest is the potential for creating a 'similarity index,' which can compare the similarity of individual surgeons using bipolar forceps, allowing for professional collaboration and refined surgical performance feedback.

ACKNOWLEDGMENTS

We acknowledge the support of the Alberta-Germany Collaborative Funds (German Canadian Centre for Innovation and Research) and the Canadian Institutes of Health Research (Grant No. 390405)-Technology Commercialization for this study. We thank all neurosurgery residents and the Division of Neurosurgery at Foothills Medical Center in Calgary, Alberta, Canada.

A.B.: investigation, methodology, software, visualization, data curation, formal analysis, writing (original draft). E.G.: methodology, software, visualization, formal analysis, writing (original draft). S.L.: conceptualization, writing (review and editing). R.S.: conceptualization, regulatory/industry compliance revision. G.R.S.: supervision, funding acquisition, writing (review and editing). All authors reviewed the manuscript.

REFERENCES

1. Choban PS, Flancbaum L. The impact of obesity on surgical outcomes: a review. *J Am Coll Surg*. 1997;185:593–603.
2. Pearse RM, Moreno RP, Bauer P, et al; European Surgical Outcomes Study (EuSOS) group for the Trials groups of the European Society of Intensive Care Medicine and the European Society of Anaesthesiology. Mortality after surgery in Europe: a 7 day cohort study. *Lancet*. 2012;380:1059–1065.
3. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. *N Engl J Med*. 2002;346:1128–1137.
4. de Vries EN, Prins HA, Crolla RM, et al; SURPASS Collaborative Group. Effect of a comprehensive surgical safety system on patient outcomes. *N Engl J Med*. 2010;363:1928–1937.
5. Birkmeyer JD, Finks JF, O'Reilly A, et al; Michigan Bariatric Surgery Collaborative. Surgical skill and complication rates after bariatric surgery. *N Engl J Med*. 2013;369:1434–1442.
6. Stulberg JJ, Huang R, Kreutzer L, et al. Association between surgeon technical skills and patient outcomes. *JAMA Surg*. 2020;155:960–968.
7. Fecso AB, Szasz P, Kerezov G, et al. The effect of technical performance on patient outcomes in surgery. *Ann Surg*. 2017;265:492–501.
8. Brodie A, Vasdev N. The future of robotic surgery. *Coll Surg Engl*. 2018;100(Supplement 7):4–13.
9. Rosen J, Hannaford B, Richards CG, et al. Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. *IEEE Trans Biomed Eng*. 2001;48:579–591.
10. Yamada H, Ahn J, Mozos OM, et al. Gait-based person identification using 3D LiDAR and long short-term memory deep networks. *Adv Robot*. 2020;34:1201–1211.
11. Zhang Z, Troje NF. View-independent person identification from human gait. *Neurocomputing*. 2005;69:250–256.
12. Gadaleta M, Merelli L, Rossi M. Human authentication from ankle motion data using convolutional neural networks. 2016 IEEE Statistical Signal Processing Workshop (SSP): IEEE, 2016:1–5.
13. Proudfoot JG, Jenkins JL, Burgoon JK, et al. More than meets the eye: how oculometric behaviors evolve over the course of automated deception detection interactions. *J Manage Inf Syst*. 2016;33:332–360.
14. Zhu H, Samtani S, Chen H, et al. Human identification for activities of daily living: a deep transfer learning approach. *J Manage Inf Syst*. 2020;37:457–483.
15. Reiley CE, Lin HC, Yuh DD, et al. Review of methods for objective surgical skill evaluation. *Surg Endosc*. 2011;25:356–366.
16. Baghdadi A, Lama S, Singh R, et al. Machine learning for segmenting and recognizing patterns of neurosurgical performance. *Sci Rep*. 2023;13:9591.
17. Sutherland G, Baghdadi A, Singh R, et al. *Machine learning-based surgical instrument characterization*. US Patent No. 17/318,975. 2021.
18. Baghdadi A, Lama S, Singh R, et al. A data-driven performance dashboard for surgical dissection. *Sci Rep*. 2021;11:1–13.
19. Sugiyama T, Lama S, Gan LS. Forces of tool-tissue interaction to assess surgical skill level. *JAMA surg*. 2018;153:234–242.
20. Sutherland GR, Zareinia K, Gan LS, et al. Bipolar forceps with force measurement. US Patent No. 13/930,361. 2021.
21. Zareinia K, Maddahi Y, Gan LS, et al. A force-sensing bipolar forceps to quantify tool–tissue interaction forces in microsurgery. *IEEE/ASME Trans Mechatron*. 2016;21:2365–2377.
22. Gan LS, Zareinia K, Lama S, et al. Quantification of forces during a neurosurgical procedure: a pilot study. *World Neurosurg*. 2015;84:537–548.
23. Sugiyama T, Gan LS, Zareinia K, et al. Tool-tissue interaction forces in brain arteriovenous malformation surgery. *World Neurosurg*. 2017;102:221–228.
24. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:770–778.
25. Fawaz HI, Lucas B, Forestier G, et al. Inceptiontime: finding alexnet for time series classification. *Data Min Knowl Discov*. 2020;34:1936–1962.
26. Wang Z, Yan W, Oates T. Time series classification from scratch with deep neural networks: A strong baseline. 2017 International joint conference on neural networks (IJCNN): IEEE, 2017:1578–1585.
27. Marchioro T, Kazlouski A, Markatos EP. *User Identification from Time Series of Fitness Data*. SECRYPT, 2021:806–811.
28. Ding H, Guo L, Zhao C, et al. RFnet: automatic gesture recognition and human identification using time series RFID signals. *Mobile Netw Appl*. 2020;25:2240–2253.
29. Seymour NE. VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. *World J Surg*. 2008;32:182–188.
30. Aïm F, Lonjon G, Hannouche D, et al. Effectiveness of virtual reality training in orthopaedic surgery. *Arthroscopy*. 2016;32:224–232.
31. Liu D, Jiang T. *Deep reinforcement learning for surgical gesture segmentation and classification. International conference on medical image computing and computer-assisted intervention*. Springer. 2018:247–255.
32. Ahmidi N, Tao L, Sefati S, et al. A dataset and benchmarks for segmentation and recognition of gestures in robotic surgery. *IEEE Trans Biomed Eng*. 2017;64:2025–2041.
33. DiPietro R, Lea C, Malpani A, et al. *Recognizing surgical activities with recurrent neural networks. International conference on medical image computing and computer-assisted intervention*. Springer. 2016:551–558.
34. Qin Y, Pedram SA, Feyzabadi S, et al. Temporal segmentation of surgical sub-tasks through deep learning with multiple data sources. 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE. 2020:371–377.
35. Wang Z, Majewicz Fey A. Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. *Int J Comput Assist Radiol Surg*. 2018;13:1959–1970.
36. Gao Y, Vedula SS, Reiley CE, et al. *Jhu-isi gesture and skill assessment working set (jigsaws): A surgical activity dataset for human motion modeling*. MICCAI workshop: M2cai: 3;2014.
37. Sheetz KH, Ibrahim AM, Nathan H, et al. Variation in surgical outcomes across networks of the highest-rated US hospitals. *JAMA Surg*. 2019;154:510–515.
38. Udyavar R, Cornwell EE, Havens JM, et al. Surgeon-driven variability in emergency general surgery outcomes: does it matter who is on call? *Surgery*. 2018;164:1109–1116.
39. Wennberg J, Gittelsohn A. Small area variations in health care delivery: a population-based health information system can guide planning and regulatory decision-making. *Science*. 1973;182:1102–1108.Cipsam la que pa con conemporem voluptuste consequatio cus ma si ut arci aut hiliquamet arum evendusanda nam aut aut reribus nonem. Et