HOKKAIDO UNIVERSITY

| | |
|---|---|
| Title | Evaluation of Laparoscopic Surgical Skill Using Surgical Instrument Motion Analysis and Machine Learning |
| Author(s) | 晏, 凌波 |
| Degree Grantor | 北海道大学 |
| Degree Name | 博士(情報科学) |
| Dissertation Number | 甲第17045号 |
| Issue Date | 2026-06-30 |
| DOI | https://doi.org/10.14943/doctoral.k17045 |
| Doc URL | https://hdl.handle.net/2115/100098 |
| Type | doctoral thesis |
| File Information | Lingbo_Yan_review.pdf, 審査の要旨 https://creativecommons.org/licenses/by/4.0/ |

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学位論文審査の要旨

博士の専攻分野の名称　博士 (情報科学)　氏名　晏　凌波

審査担当者　主　査　教　授　近野 敦
副　査　教　授　田中 孝之
副　査　客員教授　宮田 なつき
副　査　教　授　安部 崇重　（北海道大学・大学院医学研究院）

学位論文題名

Evaluation of Laparoscopic Surgical Skill Using Surgical Instrument Motion Analysis and Machine Learning
（手術器具動作解析と機械学習による腹腔鏡手術技量評価）

腹腔鏡手術は, 患者の腹部に複数個所の穴を開け, その穴にトロッカー (術具挿入口) を設置し, トロッカーから内視鏡や手術器具を挿入して行う手術である. 低侵襲で, 開腹手術と比較すると患者への負担が少なく術後の回復が早いため, 消化器外科, 泌尿器科等の領域では腹腔鏡手術 (ロボット支援手術を含む) が大多数となっている. 一方で, 術者はモニターに映し出される内視鏡映像を見ながら手術を行うことになるため, 視覚座標系と手元座標系の不一致や視野の狭さ, 奥行き感の消失等で, 術者には高度な技量が要求される. そのため、効率的かつ体系的な手術教育手法の確立が重要であるが, 従来広く用いられてきた Global Operative Assessment of Laparoscopic Skills (GOALS) や Objective Structured Assessment of Technical Skills (OSATS) 等の熟練医による主観的評価法は, 評価者間のばらつきや再現性の低さが課題として残されている. このような背景から,AI 等を活用した効率的な外科手術教育法への要求が高まっている.

本論文は以上の問題を解決するために, 光学式モーションキャプチャー装置を用いて手術器具動態の特徴量を解析し, その特徴量をもとに機械学習を行うことで, 腹腔鏡手術での手術技量を客観的, 定量的, かつ再現性良く評価するシステムを提案するもので, 全編 7 章からなる.

第 1 章は序論であり, 本研究の背景および目的を述べている.

第 2 章では, 手術技量を数値化するために構築した, 複数の手術器具の動作を光学式モーションキャプチャー装置で計測するシステムについて述べている. また, 手術技量を計測する対象であるカダバー (故人の遺志により提供されたご献体) を用いた模擬手術, 実手術, ブタの臓器を用いた模擬手術について, 実験環境の詳細およびそれぞれの研究に対する倫理審査委員会の承認について述べている.

第 3 章では, 計測した手術器具の動作の特徴量と, その算出方法を述べている. 第 4 章で述べる手術技量クラス分類, 第 5 章で述べる手術技量の数値化において, どのような特徴量を用いるかは極めて重要であり, 腎泌尿器外科医師と議論の上, 注意深く設定している.

第 4 章では, 機械学習を用いた手術技量クラス分類について述べている. モデルの精度検証は内側 (最適パラメータ探索) と外側 (精度検証) の 2 段階の入れ子状となった k 分割交差検証を繰り返す Nested k-fold cross validation を用いた. データセット分割による検証結果への影響を抑えるために, 一連の検証を 100 回繰り返す Repeated cross validation と併用している. クラス分類では, 術者

を経験手術件数で初級者, 中級者, 上級者に分類する 3 群分類と, 腹腔鏡手術技術認定取得者 (ESSQ 有) と非取得者 (ESSQ 無) に分類する 2 群分類の, 二つのクラス分類で精度を検証している. カダバーを用いた腹腔鏡下根治的左右腎摘除術 (経腹膜アプローチ) で, 提案する手術技量の計測と評価を行った. 腎摘除術を, 工程 1(前処理), 工程 2(腎血管止血・切断, 腎摘除), 工程 3(後処理) の三つに分け,46 回の模擬手術で, それぞれの工程別, および全体で手術器具動態の計測を行い, 特徴量を算出した. 算出された特徴量を入力とし, 上記の 2 群または 3 群の分類を教師信号として機械学習を行い, 出力される分類結果の精度を検証している. 3 群分類問題では, 手術全体の特徴量を用いた機械学習で分類の正解率 (accuracy) が 74〜76% と良好な結果を得た. また 2 群分類問題では, 手術全体の特徴量を用いた機械学習で分類の正解率 (accuracy) が 84〜89% となった. さらに, 説明可能 AI の一つである Shapley Additive exPlanations (SHAP) を用いて, それぞれの特徴量が分類の正解率にどの程度寄与しているかを視覚化した. 実手術に極めて近いカダバーを用いた模擬手術で, 大規模に手術技量を計測した例は国内外に見当たらず, 極めて重要な成果である.

第 5 章では, 第 4 章と同様, カダバーを用いた腹腔鏡下根治的左右腎摘除術 (経腹膜アプローチ)46 例の模擬手術のデータを用い, 熟練者 2 名による GOALS スコア (5〜25 点) の平均値を教師信号として機械学習を行い, 手術器具の動態特徴量から GOALS スコアを出力する手法を提案している. Nested and repeated k-fold cross validation で精度検証を行い, 誤差の評価指標を熟練者 2 名の GOALS スコアの平均値からの平均絶対誤差 (MAE) とし, 腎摘除術の工程 1(前処理), 工程 2(腎血管止血・切断, 腎摘除), 工程 3(後処理) のいずれにおいても, 最も精度の高い機械学習モデルでは MAE が 3 以下となり, 良好な結果を示した. 2 時間におよぶカダバーを用いた模擬手術で, 機械学習モデルを用いて精度よく手術技量を数値化できることを示したことは極めて重要な成果であり, 今後, 効率的な外科手術訓練システムへと発展することが大いに期待できる.

第 6 章では, 本研究成果の応用として, 実手術での手術手技計測, 腹腔鏡手術トレーニングボックスとブタの臓器を用いた模擬手術評価システム, 遠隔手術支援について述べている. 特に実手術での手術手技計測は世界的にも例がなく, 今後, 手術技量評価だけではなく医療事故が発生した場合の手術記録としても大いに期待でき, 非常に重要な成果であるといえる.

第 7 章は結論である.

これを要するに, 著者は実手術に極めて近いカダバーを用いた模擬腹腔鏡手術で, 術者の技量を分類または数値化しその評価理由とともに術者に提示する AI 手術技量評価システムを提案したもので, 本成果により効率的な外科手術教育が実現できるのみならず, 実手術での手術記録や手術ナビゲーションへの発展が大いに期待できることを考慮すると, 情報科学および医工学の発展に寄与するところ大である.

よって著者は, 北海道大学博士 (情報科学) の学位を授与される資格あるものと認める.