Kobe University Repository : Kernel

PDF issue: 2024-05-01

Automated Surgical Instrument Detection from Laparoscopic Gastrectomy Video Images Using an Open Source Convolutional Neural Network Platform

Yamazaki, Yuta

(Degree)
博士（医学）

(Date of Degree)
2020-09-25

(Resource Type)
doctoral thesis

(Report Number)
甲第7861号

(URL)
https://hdl.handle.net/20.500.14094/D1007861

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。

（課程博士関係）

学位論文の内容要旨

Automated Surgical Instrument Detection from Laparoscopic Gastrectomy Video Images Using an Open Source Convolutional Neural Network Platform

オープンソースのニューラルネットワークプラットフォームを用いた腹腔鏡下胃切除術における手術動画中の手術器具自動認識システムの開発

神戸大学大学院医学研究科医科学専攻
食道胃腸外科学
（指導教員：掛地　吉弘教授）

山崎　　悠太

<背景>

腹腔鏡下手術の普及に伴い生成される動画データは膨大で、手術技術の向上を目指す外科医たちがそれらを見直すことは困難である。外科教育を促進するため、腹腔鏡手術動画のコンピュータを用いた解析の開発が必要である。我々は、オープンソースの汎用物体検出ニューラルネットワークである YOLOv3 を用いて、腹腔鏡下胃切除術の手術動画から手術器具を自動認識するシステムを開発した。

<方法>

52 例の腹腔鏡下胃切除術の手術動画から 10,716 枚の画像を切り出し、訓練データセットと検証データセットを作成した。グラフィックアクセラレータを搭載したコンピュータを用いて YOLOv3 を訓練し、200,000 イテレーションの学習を行った。

システムの精度評価については、別途用意した 10 症例の手術動画を用いて、2 種類の方法で行った。まず、4 症例分の特定のシーンから一定の割合で抽出した 3,117 枚の画像をテスト用データセットとし、術者が操作する主要な 10 種類の器具について、テスト用データセット中での認識精度を算出した。

さらに、10 症例分の動画全体を本システムで解析して、すべての手術器具の検出頻度を時系列で図示したヒートマップを症例ごとに作成した。そのヒートマップを 3 人の外科医が評価し、ヒートマップが手術で実際に使用された手術器具と対応しているか、再建方法や出血イベント等を正しく表しているか、を評価した。

<結果>

テスト用データセットを用いた精度評価では、適合率は 0.87、感度は 0.83 であった。手術器具検出頻度を表したヒートマップは、使用された手術器具の種類、胃切離・十二指腸切離のタイミング、再建方法の種類、出血イベントの有無等の手術の特徴を 90%以上の精度で正しく表していた。

<考察>

腹腔鏡下手術画像における画像認識についての既存研究の多くは、腹腔鏡下胆嚢摘出術を対象としており、それらの研究ではオープンソースの大規模な画像データセットを用いた精度評価が行われている。一方で、腹腔鏡下胃切除術はより多くの手術器具や術野展開を要する術式であり、腹腔鏡下胃切除術を対象とした既存の画像データセットは存在しない。公開された画像データセットを用いずに精度評価を行う場合は、明確な基準をもって作成した膨大な画像データセットなくしては、画像の選択バイアス等の問題から、使用時間の極めて短い手術器具や、術野に不意に映り込むような小さな手術器具などについての認識精度を公正に判定することは難しい。上記から本研究では、2 種類の方法でシステムの精度評

価を行った。

使用頻度の多い主要な10種の手術器具については、テスト用画像データセットを用いて評価し、既存研究と同様の高い認識精度を確認した。一方で、画像データセットを用いた評価では公正な評価が難しいと思われた手術器具（はさみ、クリップなど）を含んだ全手術器具を対象としたヒートマップを用いた精度評価では、使用された手術器具の種類だけでなく、手術手技のタイミングや再建方法・出血イベント等までも、外科医による解釈によりヒートマップから読み取ることができ、これを以て精度評価とした。後者の評価法は、実際の手術データとの整合性の確認や、再建方法や出血イベントの解釈が必要であり、胃外科の実臨床に携わる外科医の視点無くしては成し得ないような本研究独自の手法である。

ディープラーニングを用いた画像認識の手法としては、本研究で用いた物体検出以外に、画像をピクセルレベルでクラス分類するセマンティックセグメンテーションがあり、いずれも学習・実装のために画像データセットを作成する必要がある。物体検出は手術器具の使用頻度解析や鉗子の軌道解析に有用であると考えられ、画像中の物体をバウンディングボックスで囲んで注釈付けをするという比較的簡易な作業で、データセットを作成可能である。それに対して、セマンティックセグメンテーションはより詳細な画像認識が可能で、術中ナビゲーション等に有用であると考えられるが、データセット作成には多大な労力と時間を要する。よって目的に応じて、適切なアルゴリズムを選択することが重要である。

本研究のリミテーションとしては、精度評価にオープンソースの画像データセットを用いていないこと、テスト用データセットのサンプルサイズが小さいこと、ヒートマップからの出血イベントの解釈はヒートマップ中の淡い色合いの読み取りによって為されておりややあいまいであること、などが挙げられる。

＜結論＞

オープンソースの汎用物体検出ニューラルネットワークであるYOLOv3を用いて、腹腔鏡下胃切除術の手術動画における、手術器具自動認識システムを開発した。本システムの0出力結果を使用して作成したヒートマップを見ることにより、外科医は手術動画を見ることなく、手術器具の使用パターンや術中のワークフローを想起することができ、膨大な手術動画の見直しをより効率的にできる可能性がある。

神戸大学大学院医学(系)研究科（博士課程）

論文審査の結果の要旨

受付番号	甲 第 2993 号	氏 名	山崎 悠太
論文題目 Title of Dissertation	Automated Surgical Instrument Detection from Laparoscopic Gastrectomy Video Images Using an Open Source Convolutional Neural Network Platform オープンソースのニューラルネットワークプラットフォームを用いた腹腔鏡下胃切除術における手術動画中の手術器具自動認識システムの開発		
審査委員 Examiner	主 査 Chief Examiner 福本巧 副 査 Vice-examiner 関口、裕之 副 査 Vice-examiner 佐々木 良平		

（要旨は１，０００字～２，０００字程度）

【背景】
腹腔鏡下手術の普及に伴い生成される動画データは膨大で、手術技術の向上を目指す外科医たちがそれらを見直すことは困難である。外科教育を促進するため、腹腔鏡手術動画のコンピュータを用いた解析の開発が必要である。我々は、オープンソースの汎用物体検出ニューラルネットワークであるYOLOv3を用いて、腹腔鏡下胃切除術の手術動画から手術器具を自動認識するシステムを開発した。

【方法】
52例の腹腔鏡下胃切除術の手術動画から10,716枚の画像を切り出し、訓練データセットと検証データセットを作成した。グラフィックアクセラレータを搭載したコンピュータを用いてYOLOv3を訓練し、200,000イテレーションの学習を行った。
システムの精度評価については、別途用意した10症例の手術動画を用いて、2種類の方法で行った。まず、4症例分の特定のシーンから一定の割合で抽出した3,117枚の画像をテスト用データセットとし、術者が操作する主要な10種類の器具について、テスト用データセット中での認識精度を算出した。
さらに、10症例分の動画全体を本システムで解析して、すべての手術器具の検出頻度を時系列で図示したヒートマップを症例ごとに作成した。そのヒートマップを3人の外科医が評価し、ヒートマップが手術で実際に使用された手術器具と対応しているか、再建方法や出血イベント等を正しく表しているか、を評価した。

【結果】
テスト用データセットを用いた精度評価では、適合率は0.87、感度は0.83であった。手術器具検出頻度を表したヒートマップは、使用された手術器具の種類、胃切離・十二指腸切離のタイミング、再建方法の種類、出血イベントの有無等の手術の特徴を90%以上の精度で正しく表していた。

【考察】
腹腔鏡下手術画像における画像認識についての既存研究の多くは、腹腔鏡下胆嚢摘出術を対象としており、それらの研究ではオープンソースの大規模な画像データセットを用いた精度評価が行われている。一方で、腹腔鏡下胃切除術はより多くの手術器具や術野展開を要する術式であり、腹腔鏡下胃切除術を対象とした既存の画像データセットは存在しない。公開された画像データセットを用いずに精度評価を行う場合は、明確な基準をもって作成した膨大な画像データセットなくしては、画像の選択バイアス等の問題から、使用時間の極めて短い手術器具や、術野に不意に映り込むような小さな手術器具などについての認識精度を公正に判定することは難しい。上記から本研究では、2種類の方法でシステムの精度評価を行った。

使用頻度の多い主要な 10 種の手術器具については、テスト用画像データセットを用いて評価し、既存研究と同様の高い認識精度を確認した。一方で、画像データセットを用いた評価では公正な評価が難しいと思われた手術器具（はさみ、クリップなど）を含んだ全手術器具を対象としたヒートマップを用いた精度評価では、使用された手術器具の種類だけでなく、手術手技のタイミングや再建方法・出血イベント等までも、外科医による解釈によりヒートマップから読み取ることができ、これを以て精度評価とした。後者の評価法は、実際の手術データとの整合性の確認や、再建方法や出血イベントの解釈が必要であり、胃外科の実臨床に携わる外科医の視点無くしては成し得ないような本研究独自の手法である。

ディープラーニングを用いた画像認識の手法としては、本研究で用いた物体検出以外に、画像をピクセルレベルでクラス分類するセマンティックセグメンテーションがあり、いずれも学習・実装のために画像データセットを作成する必要がある。物体検出は手術器具の使用頻度解析や鉗子の軌道解析に有用であると考えられ、画像中の物体をバウンディングボックスで囲んで注釈付けをするという比較的簡易な作業で、データセットを作成可能である。それに対して、セマンティックセグメンテーションはより詳細な画像認識が可能で、術中ナビゲーション等に有用であると考えられるが、データセット作成には多大な労力と時間を要する。よって目的に応じて、適切なアルゴリズムを選択することが重要である。

本研究のリミテーションとしては、精度評価にオープンソースの画像データセットを用いていないこと、テスト用データセットのサンプルサイズが小さいこと、ヒートマップからの出血イベントの解釈はヒートマップ中の淡い色合いの読み取りによって為されておりややあいまいであること、などが挙げられる。

【結論】

オープンソースの汎用物体検出ニューラルネットワークである YOLOv3 を用いて、腹腔鏡下胃切除術の手術動画における、手術器具自動認識システムを開発した。本システムの０出力結果を使用して作成したヒートマップを見ることにより、外科医は手術動画を見ることなく、手術器具の使用パターンや術中のワークフローを想起することができ、膨大な手術動画の見直しをより効率的にできる可能性がある。

本研究は腹腔鏡下手術動画において人工知能を用いて手術器具を自動認識するシステムの開発について研究したものであるが、従来ほとんど検討されていなかった胃切除術動画における物体検出モデルの応用とそのシステムの実用化について重要な知見を得たものとして価値ある集積であると認める。よって、本研究者は、博士（医学）の学位を得る資格があると認める。