

厚生労働科学研究費補助金  
(政策科学総合研究事業(臨床研究等 ICT 基盤構築・人工知能実装研究事業))

分担研究報告書

ロボット型内視鏡操作支援システムの AI による高度化と  
各種医療機器統合インターフェースとしての展開に関する研究

研究分担者 柴田 裕一郎 長崎大学大学院工学研究科 教授

研究要旨 内視鏡から得られる画像データを解析し、執刀医にとって望ましい画角となるように内視鏡操作ロボットを自動制御するメカニズムの要素技術として、胆嚢摘出手術を例に、入力画像の角画素を臓器部位に従って 4 クラスに分類するセグメンテーションシステムを畳み込みニューラルネットワークを用いて構築した。深層学習技術を医療分野に応用する場合、利用可能な教師付きデータセットの量が少ないという問題がある。そこで過学習による品質低下の抑制を図るため、ネットワークの構造を工夫することでパラメータ数を削減する手法を提案した。実験の結果、672×528 ピクセルの画像に対し、最大 82.9%の正答率をフレームレート 55.40 fps で達成できることを確認した。

共同研究者

眞邊 泰斗、友永 航生(長崎大学)

A. 研究目的

本研究課題「ロボット型内視鏡操作支援システムの AI による高度化と各種医療機器統合インターフェースとしての展開」の目的を達成するためには、内視鏡から得られる画像データを解析し、執刀医にとって望ましい画角となるように内視鏡操作ロボットを制御するメカニズムの実装が必要である。このためには、内視鏡画像の中の臓器部位を識別し、それらの配置や画角内に占める面積割合などをリアルタイムに算出することが必要となる。このような処理には、深層学習に基づく畳み込みニューラルネットワークを用いた AI による画像セグメンテーション技術の応用が有望である。しかし、AI 技術を医療分野に応用する場合、学習用に利用可能な教師付きデータセットの量が極めて少ないという本質的な問題がある。このため、画像セグメンテーションに有効なニューラルネットワークをそのま

ま応用しようとしても、過学習の問題が生じ望ましい学習結果を得ることができない。そこで、既存のネットワークを参考にしつつ、小規模なデータセットでも学習が可能な新しいニューラルネットワークの構成を提案し、その効果と課題を明らかにすることを研究目的とした。

B. 研究方法

画像セグメンテーション分野での効果が高いことで知られる畳み込みニューラルネットワークの一種である U-Net をベースとし、構造を単純化したネットワークを提案した。U-Net 全体が 1 つの大きなネットワークとして設計されているのに対し、本提案ネットワークは縮退用の小規模サブネットワークと特徴マップ連結用の小規模サブネットワークを再帰的に用いる構造とした。このことにより、少ないパラメータで学習を行い、過学習を抑制することを可能とし

ている。また、これらのサブネットワークは、複数の異なる解像度をもつ画像に対して処理を行う構成とすることで、特定のサイズにとらわれずに、より普遍的な特徴を抽出するよう工夫している。このことも過学習の抑制につながりうる。

提案ネットワークの学習には Python 上で作動するニューラルネットワーク向けフレームワークである Chainer 1.17.0 を用い、また学習高速化のために NVIDIA 社の GPU 向けコンピューティングフレームワーク CUDA を用いた。学習および評価に必要な各種処理には、数値計算ライブラリの Numpy 1.8.2、画像処理ライブラリの scikit\_image 0.12.3 および Pillow 4.2.1 を用いた。

学習に用いる訓練用データセットと、分類の評価に用いる評価データセットはそれぞれ 23 枚、8 枚であり、画像サイズは 672 × 528 ピクセルのものを用了。これらは実際の胆嚢摘出手術の記録映像から抽出した静止画像であり、これらを胆嚢、胆嚢管、総胆管、その他の 4 つのクラスに医師が手動でラベリング・アノテーションを行ったものである。学習にはミニバッチ学習を採用した。訓練データから 256 × 256 ピクセルの小領域をランダムに切り出し、水平・垂直方向にランダムで反転させた後、これらを 4 組ずつ束ねることで生成した。損失関数には Softmax Cross Entropy を用い、最適化関数には Adam を用いた。提案ネットワークの構成については、パラメータを変更することで、ノード数が 12、16、30 の 3 種類のネットワーク構成を評価し比較した。

(倫理面への配慮)

学習用のデータについては個人が特定され得ない手術画像を用いた。

### C. 研究結果

まず、セグメンテーションの正答率(全画素のセグメンテーション結果を教師データと比較したとき、全画素数に占める正しいクラスに分類された画素数の割合)の評価を行った。その結果は、ノード数が 12 のとき 79.7%、ノード数が 16 のとき 82.9%、

ノード数が 20 のとき 81.7%であった。

次に、推論の処理時間について評価を行った。実験環境として、CPU に Intel 社 Core i7-5930K 3.50GHz、GPU に Nvidia 社 GeForce GTX 1080 を用いた。それぞれ 600 枚の画像を処理した時の平均フレームレート (fps: frames per second) を測定したところ、ノード数が 12 のとき 56.71 fps、ノード数が 16 のとき 55.40 fps、ノード数が 20 のとき 55.17 fps であった。

また、これらのセグメンテーション結果に基づき、本課題研究分担者である中央大学の諸麥准教授らが作成した内視鏡操作ロボットを操作できるよう、ロボット接続のためのインタフェースを非同期シリアル通信を用いて実装した。接続実験の結果、ニューラルネットワークによってロボットを正常に制御できることを確認した。

### D. 考察

セグメンテーションについては概ね 80% 程度の良好な正答率が得られている。一方、ノード数が 20 のときの正答率は、ノード数が 16 のときよりもわずかに低下しており、この結果は過学習の影響が反映されたものと解釈できる。クラスラベリングの結果画像についても、臓器の位置を概ね妥当に捉えることができていることが確認できた。学習用データセットが少ない中、過学習による品質劣化をある程度抑制することができており、提案ネットワークの構造上の工夫により一定程度の成果を得ることができたと判断できる。

また、ニューラルネットワークの構成をコンパクトにしたことにより、推論処理に要する時間の短縮にもつながっていることが、性能評価の結果から伺える。GPU による計算アクセラレーションの効果ともあいまって、実験に用いたいずれの構造の場合でも 55 fps を上回るリアルタイム性が確認できた。この処理性能は、ロボット制御の観点からも十分なものであると判断できる。内視鏡自動操作ロボットの接続テストについても正常な動作確認が行えたことから、今後は学習用データセットを追加することによりセグメンテーションの確度を向上し

ていくことが今後の主たる課題として挙げられる。

CANREXI Workshop (2018.11)

#### E. 結論

胆嚢摘出手術における内視鏡の自動操作を可能とする要素技術のひとつとして、入力された画像の各画素を胆嚢、胆嚢管、総胆管、その他の4クラスに分類する画像セグメンテーションシステムを畳み込みニューラルネットワークを用いて構築した。少ないデータセットで過学習の影響を抑制するために、医療画像セグメンテーション用に考案されたU-Netをベースとしつつ、小さなサブネットワークを再帰的に適用する構成を提案し、パラメータ数を削減した。評価の結果、672×528ピクセルの画像に対し、最大82.9%の正答率をフレームレート55.40fpsで達成できることを明らかにした。

#### H. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得  
なし
2. 実用新案登録  
なし
3. その他  
なし

#### F. 健康危険情報

なし

#### G. 研究発表

##### 1. 論文発表

T. Manabe, Y. Shibata, K. Oguri, FPGA Implementation of a Real-Time Super-Resolution System Using Flips and an RNS-Based CNN, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E101-A, No.12, pp.2280-2289 (2018.12)

H. Egawa, Y. Shibata, Storing and Compressing Video into Neural Networks by Overfitting, Advances in Intelligent Systems and Computing, Vol.772, pp.615-626 (2018.7)

##### 2. 学会発表

Y. Shibata, Near-I/O reconfigurable computing for medical engineering, International Symposium on Computing and Networking,