

画像計測と教師付きデータ生成のための 作業支援システムの提案

榎本 洗一郎^{†1} 石井 彩室^{†1} 戸田 真志^{†2} 川崎 琢真^{†3} 清水 洋平^{†3}

概要：深層学習の技術革新により、より大量で正確なアノテーションが付与された教師付きデータセットの重要性が増している。一方で、データが蓄積されていない分野では、新たに膨大な教師付けの作業を行う必要があり、大きな負担となる。そこで本稿では、新たに通常の画像計測の作業のみで、正確なアノテーションの付与が保証される教師付きデータ生成支援システムを提案する。また実際に提案ツールを用いてセマンティックセグメンテーションの手法で代表的な U-Net を用いて、得られたデータセットを学習し、精度検証し、その有効性を示す。

1. はじめに

近年の深層学習の急速な進歩により、大量かつ正確な教師付きデータセットの重要性が増している。データセットは、機械学習を用いて学習させたい課題に応じて、元となる画像や音声情報などを整理し、アノテーションを付けるなどの教師付けをして作成される。一方で、様々な分野で深層学習の技術導入を検討する際には、通常の作業に加えて、教師付きデータの作成が求められ、専用のソフトウェアも提供されている[1]が、大きな負担であり障害となっている。そこで本稿では、新たにタッチ教示を用いた画像計測支援と同時に、教師付きデータを生成するシステムを提案する。提案システムの概略を図 1 に示す。

著者らは、計測対象を含む画像に対して計測領域を"なぞる"というタッチ教示により "ハカル"を実現する画像計測支援ツール「TouchDeMeasure」を提案している [2]。

TouchDeMeasure では、通常のデジタル画像や顕微鏡画像に限らず空撮画像など様々な画像を想定している。またユーザに求められる操作は、正確なクリック・ドラッグ操作ではなく、対象領域内を"なぞる" (ドラッグ・タッチ教示) 操作であり、この情報に基づき計測対象に適した計測項目をシステムが自動計測するものである。

本稿では、通常業務の中で行われている画像計測の結果は、その目的に対して十分な精度が担保されているものであるところに注目し、これまでに提案している画像計測支援ツール「TouchDeMeasure」に教師付きデータ生成の機能を組み込み、新たな教師付きデータ生成ツールを提案する。また実際に提案ツールによって得られた教師付きデータを用いて、セマンティックセグメンテーションの手法で代表的な U-Net[3]を用いて、その有効性について示す。

2. システムの提案

2.1 画像計測

一般的に画像計測において、動植物のサンプリング調査

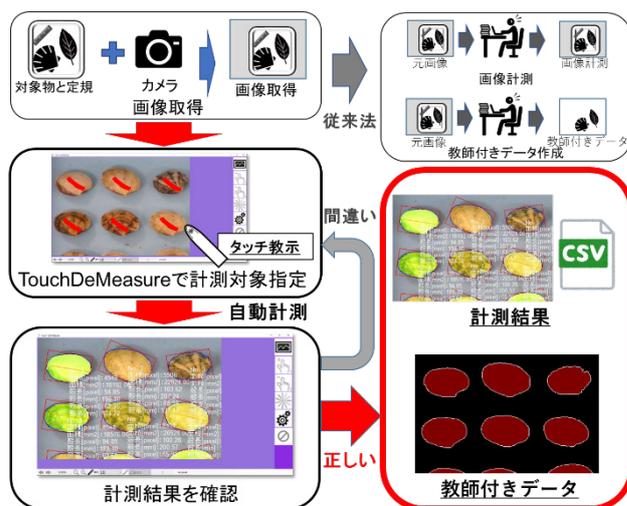


図 1 従来法と提案する教師付きデータ生成支援システムの概略図

や顕微鏡画像を用いる場合、デジタルカメラ等で計測対象と定規などのスケールを撮影した画像から測定される。しかしこれらの測定は、対象に適した複雑な計測項目を大量に行うことも多く、大きな負担となっている。特に面積や周囲長等を計測する際は、対象領域の端を正確に指定する必要があり、最大のボトルネックとなっている。

画像計測は、対象とする平面を記録し、計測することから、様々な測定との親和性が高い。ユーザの目視による画像計測（従来法）ではマウス操作が基本であり一般的な計測項目を例とすると、数は対象のクリック操作、長さは対象となる2点以上のクリック操作、面積は対象領域の塗りつぶし（ドラッグ操作）などとなる。これらの作業は全て正確に行う必要があり、クリック・ドラッグ操作のズレは計測誤差につながる。

ユーザがデジタル画像からモノを計測する作業は2つに大別すると、計測対象（領域）か否かの選別、計測対象に適した項目について計測、することである。前者においては、対象画像中に含まれるモノが全て計測対象となるとは限らず、ゴミやノイズ、または計測目的に適さないものも含まれる。また撮影環境は、屋内環境下で撮影された画

^{†1} 滋賀県立大学

^{†2} 熊本大学

^{†3} (地独) 北海道立総合研究機構

像に限らず、屋外環境や海中、空撮、顕微鏡など様々な状況が想定される。このため、対象画像から対象物体を自動で計測する場合、それぞれの撮影環境を考慮し、複数のモノから計測対象領域のみを抽出する必要がある。後者は、対象により計測項目が異なるが、それぞれの計測項目や測り方は各々定義されており、計測者は計測に必要な専門的な知識が求められる。

このため、画像計測に用いられたデータは、少なくとも画像計測が必要な目的に対して、十分な精度が保証されているものであるが、本来の目的でしか活用されていないのが現状である。

2.2 関連研究

機械学習の研究や実用化のために、いくつものデータセット作成のためのツールが提案されている。

Microsoft 社は、VoTT(Visual Object Tagging Tool)と呼ばれる画像中に、タグを埋め込むことのできるツールを提供している[1]。これは一般物体認識や物体検出などのために、対象画像にアノテーションを付与するものであり、作業を効率行うための機能が実装されている。また同様に LabelImg[4]などのソフトウェアもあるが、それぞれアノテーションの作業支援機能を備え、使用する機械学習に適したデータセット形式で保存することが可能である。

一方で、教師付けデータセットの作成には、アノテーションの付与の作業をすることが必須である(図 1)。このため従来の業務に加えて作業が発生する場合、従事者の負担が単純に増加するものである。

2.3 総合支援システムの提案

著者らは、計測対象を含む画像に対して計測領域を"なぞる"というタッチ教示により計測を実現する画像計測支援ツール「TouchDeMeasure」を開発している[2]。これに教師付きデータの生成機能を新たに追加し、画像計測の結果を教師付きデータとして活用可能なシステムを提案する。

提案システムの概要を図 2 に示す。まずユーザは対象画像および計測対象に応じた計測モードを選択する(図 2)。次に手入力またはスケールマーカーにより実スケールを算出する。ここでユーザが計測対象である前景領域および対象外の背景領域の一部をタッチ教示により指定する[5]。1 回のタッチ教示は 1 つ計測対象に対応し、

- 計測対象(領域)か否か
- 計測対象の数
- 計測対象(領域)の一部の情報

を与えることとなる。つまりユーザのタッチ教示は、計測対象か否かをユーザ自身が情報として与えると同時に、システムが計測対象の領域抽出を行う上で必要な領域の情報を提供するものである。ユーザのタッチ教示に基づき、システムは多クラス対応型領域拡張法[2]により計測対象領域を抽出する。そしてユーザによって選択された計測モードに基づき、計測が行われる。得られた結果をユーザが確

認し、計測結果に不備があればもう一度領域の指定・計測を行い、不備がなければ計測結果を保存する。このとき提案システムは、通常の画像計測の結果である計測結果の画像、画像および実スケールに変換された計測結果が記録された CSV ファイルと共に、元画像ファイル、セマンティックセグメンテーションなどに活用できる領域抽出結果画像のファイル、計測モードで指定された計測対象名を記録する。

すでに、アサリとホタテガイ、ナマコの計測に対応しており、計測対象ごとに計測モードを選択する際に、提案ツールにより計測された対象物名のアノテーションも付与することができる。

教師付きデータとして記録されるのは、領域拡張法によって抽出された結果であり、この結果は従来の画像計測の目的に対して十分な精度であるとユーザに確認されることで保証される。そしてユーザは、画像計測に不十分な精度や間違いを含む場合は、対象を除外または再計測を行うため、記録される教師付きデータの品質は保たれる。また本稿では、領域抽出結果のみを対象として紹介するが、対象領域の教師付けされているため、矩形で対象領域をタグ付けし物体検出のためのデータなども容易に拡張可能である。

実装した提案ツールの外観を図 2 に示す。開発環境は Windows10、Microsoft Visual Studio C#であり、動作環境はタッチスクリーンおよびマウス操作に対応した 64bit の Windows デスクトップアプリケーションである。基本機能として、画像の拡大・縮小表示、戻る・やり直し、手動およびマーカーによる自動スケール変換、タッチ教示の線幅指定、計測および CSV ファイルへの表示・記録の項目指定に対応している(図 2)。

3. 実験

3.1 実験方法

提案ツールで得られた教師付きデータの検証のために、U-Net を用いた領域抽出精度の検証を行った。

実験には、従来の TouchDeMeasure により水産業におけるアサリのサンプリング調査のために撮影された 24 枚の画像を使用し、画像中には計 1128 個体のアサリが含まれている。対象画像の例を図 4 に示す。またスケール自動検出用の AR マーカーが貼られた定規も含まれている。この画像に対して、提案ツールにより画像計測を行い、領域抽出された結果を教師付きデータとして得る。得られたデータを 256×256 pixel に一部重複を含めて切り出した画像を 7062 枚生成し、くわえて上下反転、拡大縮小、コントラスト補正をし、112,992 枚のデータセットを構築した。この内 96,043 枚を学習データ、16,949 枚をテストデータとして、

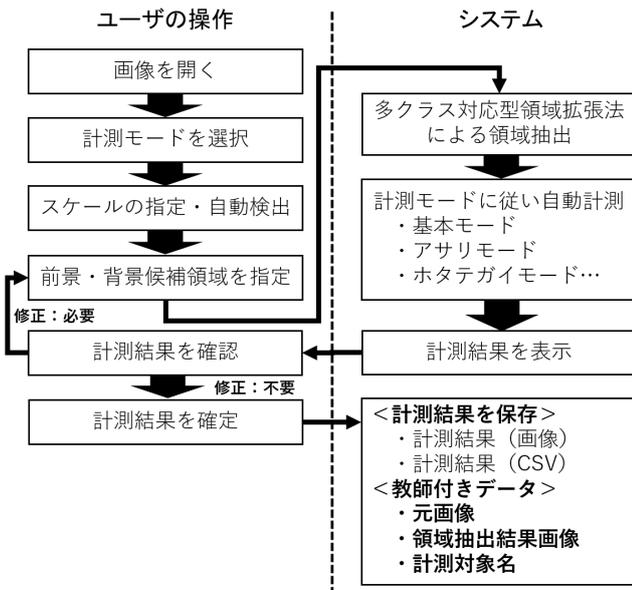


図 2 画像計測と教師付きデータ生成支援の処理の流れ



図 3 提案システムの外観と各部説明

精度を検証した。なお、学習に使用した計算機は CPU:Core i7 3.6GHz, メモリ:64GB, GPU:NVIDIA Geforce RTX2060 であり、U-Net は Epoch 数 200, Batch サイズを 32 とした。

3.2 実験結果と考察

提案ツールによる画像計測の結果例を図 5, U-Net による領域抽出の実験結果例を図 6, U-Net による学習結果を図 7 に示す。

図 5 は、従来の画像計測支援ツール「TouchDeMeasure」の結果[2]となるが、対象のアサリに対応した計測項目（殻長と殻高）が計測されており、抽出された領域も良好である。ここで得られた領域を教師付きデータとして、U-Net により学習された識別器を用いて、アサリの領域を抽出した結果も、学習精度も良好であった（図 6, 図 7）。

実験結果より、画像計測で求められる計測精度を満たした領域抽出結果を活用することで、セマンティックセグメンテーションの問題を深層学習に解くために十分なアノテーションの精度であることが示される。また特に今回のように、水産業のような利用範囲が限定されるような分野においても、通常の業務範囲内で新たなデータセットの作成

可能性を示したことで、様々な分野への応用が期待される。



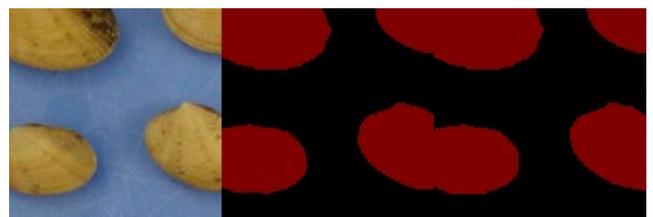
図 4 対象画像例



図 5 提案ツールによる画像計測の結果例



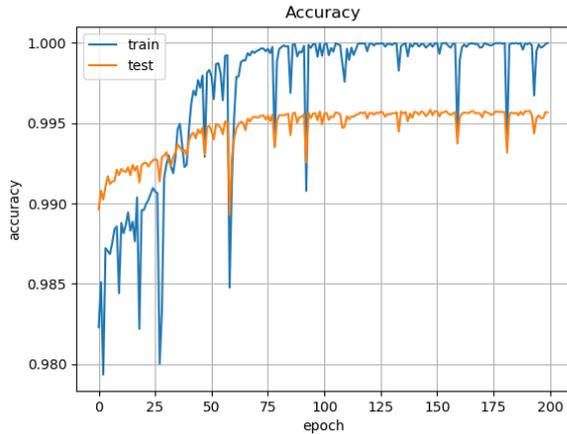
(a)



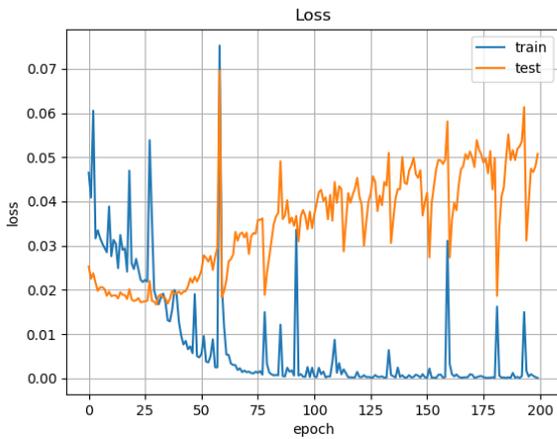
(b)

図 6 U-Net による実験結果例。

左：元画像 中：予測画像 右：正解画像



(a) 学習精度の推移



(b) 損失推移

図 7 U-Net の学習結果

- [4] “tzutalin: LabelImg”, <https://github.com/tzutalin/labelImg/>, (参照 2019-12-20).
- [5] 清野達也, 林貴宏, 尾内理紀夫, 三條正裕, 森正弥: 改良領域拡張法による高速画像切り抜き手法の提案と評価, 情報処理学会論文誌, Vol.50, No.12, pp.3233--3249 (2009).

4. おわりに

画像計測の作業に注目し, 新たに画像計測支援ツールに教師付きデータの生成機能を追加することで, 教師付きデータ作成の作業を行わずにデータを蓄積可能なシステムを提案した. また提案システムによって得られたデータセットにより深層学習による領域抽出の実験結果を示し, その有効性を示した.

今後は, 実際の教師付けの作業時間や工程などとの比較と共に, さらなる機能拡充を目指す予定である.

参考文献

- [1] “Microsoft: VoTT(Visual Object Tagging Tool)”, <https://github.com/Microsoft/VoTT>, (参照 2019-12-20).
- [2] 榎本洗一郎, 戸田真志, 清水洋平, 宮崎義弘, 吉田真也: 水産資源管理のためのユーザ支援型画像計測システムの提案, 動的画像処理実用化ワークショップ (DIA2015), 4 pages in CD-ROM, (2015).
- [3] Ronneberger, O., P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." Medical Image Computing and Computer-Assisted Intervention (MICCAI). Vol. 9351, pp. 234-241 (2015).