

令和 5 年 6 月 8 日現在

機関番号：12608

研究種目：若手研究

研究期間：2020～2022

課題番号：20K19837

研究課題名（和文）微分可能クラスタリングによる教師なし画像セグメンテーションの深層学習に関する研究

研究課題名（英文）A Study on Deep Learning of Unsupervised Image Segmentation by Differentiable Clustering

研究代表者

金崎 朝子（Kanezaki, Asako）

東京工業大学・情報理工学院・准教授

研究者番号：00738073

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究では教師データを必要としない教師なし深層学習による画像セグメンテーションを開発した。デファクトスタンダードな教師なし画像セグメンテーション手法であるグラフカット、および近年開発された深層学習ベースの従来手法と比較し、複数のベンチマークデータセットにおいて提案手法の高い有効性を示した。本研究成果は画像処理分野のトップジャーナルIEEE TIP（IF：9.34）に採択された。さらに、電気通信普及財団 テレコムシステム技術学生賞、およびIEEE Signal Processing Society (SPS) Japan Student Journal Paper Awardを受賞した。

研究成果の学術的意義や社会的意義

本研究成果を再現するソースコードをオープンソースとしてGitHubに公開しており、既に600弱のスター数を獲得している。当該ソースコードは、世界中の様々な大学や研究機関において、特に医用画像処理分野で広く利用されている。さらに、「教師なし画像セグメンテーションのベーシックな手法と深層学習ベースの手法の紹介」という論文タイトルで、日本医用画像工学会(JAMIT)誌「MEDICAL IMAGING TECHNOLOGY 39(4)」の特集論文を寄稿した。

研究成果の概要（英文）：In this study, we developed an unsupervised deep learning image segmentation method that does not require supervised data. Compared to graph cuts, the de facto standard unsupervised image segmentation method, and recently developed deep learning-based conventional methods, the proposed method is highly effective on several benchmark datasets. The research results have been accepted for publication in IEEE TIP (IF: 9.34), a top journal in the field of image processing. Furthermore, we received the Telecommunications Advancement Foundation Telecom System Technology Student Award and IEEE Signal Processing Society (SPS) Japan Student Journal Paper Award.

研究分野：知能情報処理

キーワード：画像処理 深層学習 教師なし学習 画像セグメンテーション

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

昨今の機械学習分野の発展は目覚ましく、画像認識をはじめとする様々な課題で深層学習が使われるようになった。従来は人間が設計した特徴量を抽出していた部分を深層学習のレイヤーに置き換え End-to-End な学習を行うことで、高精度なパターン認識が可能となる。さらに、パターン認識のみにとどまらず、画像や 3D モデルの生成、対応点探索、幾何情報の推定といった様々なタスクにおいて深層学習による飛躍的な高精度化が実現されてきている。ある課題を遂行するために深層学習を導入するためには、「微分可能であること」が鍵となる。深層学習は一般的に連鎖律 (Chain rule) に基づく誤差逆伝播法を用いて、定義された損失を減少させるようなパラメータの勾配を求め、その勾配方向へと少しずつ変化させることでパラメータを学習する。このような誤差逆伝播法を用いるためには、すべての処理が微分可能な処理の連鎖で構成される必要がある。

2. 研究の目的

本課題では、「微分可能な処理の連鎖で画像セグメンテーションを遂行するアルゴリズムを提案し、深層学習を行うことで、従来よりも高精度な教師なし画像セグメンテーションが実現できるか」という問いを解明する。昨今、「車」、「人」、「道路」等の意味的なまとまりを持った画像領域を分割しラベルを付与するセマンティックセグメンテーションにおいては、深層学習ベースのアルゴリズムが多く提案されている。これらのほとんどが教師あり学習であり、大量の学習データを必要とする上に、学習データにない物体はセグメンテーションできないという問題がある。これに対し、教師なし画像セグメンテーションとは教師データを必要とせず、与えられたテスト画像を任意の個数の領域に分割する課題である。事前知識を活用できないという点において教師あり学習よりも難易度が高く、申請者の知る限り、教師なし画像セグメンテーションに深層学習を適用した例はまだない。従来の教師なし画像セグメンテーションを行う手法の一つに、画素の特徴量に基づくクラスタリングが存在する。たとえば有名な k -means アルゴリズムは、特徴空間上のランダムに選択した k 点をクラスタ中心とし、最も距離の近いクラスタへの割り当てとクラスタ中心座標の更新とを交互に行うアルゴリズムであるが、各点のクラスタへの割り当てというプロセスが微分可能ではない。そこで本課題では、全結合層とバッチ正規化、そしてソフトマックス関数という微分可能な処理の組み合わせでクラスタ割当処理を構成し、微分可能クラスタリングを実現する。これにより、画素の特徴抽出とクラスタリングを全てつないだ End-to-End な深層学習を行うことが可能になる。本申請課題では、従来の画像セグメンテーションのプロセスに照らし合わせて提案手法を説明し、その妥当性・優位性を理論的および実験的に明らかにする。

3. 研究の方法

提案する微分可能クラスタリングは、 D 次元特徴空間の各軸上の無限遠点をクラスタ中心とみなし、各軸への射影距離が最小となるクラスタへと各点を割り当てることで行う。この操作を全結合層とバッチ正規化、そしてソフトマックス関数という微分可能

な処理で構成し,前段の畳み込み層による画素特徴抽出と連結して,誤差逆伝播法により End-to-End な学習を行う.一連の処理内容は,大雑把に言えば,デファクトスタンダードな従来手法である k -means クラスタリングを近似的に深層学習で実現し,特徴量学習と同時に行うことを可能にするものである.提案手法は,以下の損失を最小化する.

$$L = \sum_{n=1}^N \sum_{d=1}^D -\delta(d - c_n) \ln v_{n,d} + \left| \frac{dv_{n,d}}{dx} \right| + \left| \frac{dv_{n,d}}{dy} \right|$$

ただし N は画素数, $v_{n,d}$ は n 番目の画素に対するネットワークの d 次元目の出力, c_n は割当クラスタのクラスタ番号, x と y はそれぞれ画像の横軸と縦軸を表す.この損失を最小化することで,互いに特徴量が類似している画素を空間的に滑らかに同じクラスタへと集合させることができる.

4. 研究成果

提案手法は入力画像に対し,特徴抽出モジュールと画素クラスタリングモジュールを End-to-End に学習することで,双方を最適化する.(特徴学習を行わない)デファクトスタンダードな教師なし画像セグメンテーション手法であるグラフカット,および近年開発された深層学習ベースの従来手法と比較し,複数のベンチマークデータセットにおいて提案手法の高い有効性を示した.また,scribble と呼ばれるユーザインプットにより少数の画素の正解ラベルを手がかりとして与えるセグメンテーションタスクにおいても,提案手法が非常に高い精度を出すことを示した.さらに,二次元画像だけでなく動画データへのセグメンテーションへ応用し,様々なアプリケーションへと発展させた.本研究成果は画像処理分野のトップジャーナルである IEEE Transactions on Image Processing (IF: 9.34) に採択された.さらに,電気通信普及財団 テレコムシステム技術学生賞,および IEEE Signal Processing Society (SPS) Japan Student Journal Paper Award を受賞した.

本研究成果を再現するソースコードをオープンソースとして GitHub に公開しており,既に 600 弱のスター数を獲得している.当該ソースコードは,世界中の様々な大学や研究機関において,特に医用画像処理分野で広く利用されている.さらに,「教師なし画像セグメンテーションのベーシックな手法と深層学習ベースの手法の紹介」という論文タイトルで,日本医用画像工学会(JAMIT)誌「MEDICAL IMAGING TECHNOLOGY 39(4)」の特集論文を寄稿した.最後に,提案手法の出力結果であるセグメンテーション画像を入力とした画像のシーン識別手法を開発し,コンピュータビジョン分野における国内最大規模の会議である MIRU2022 にて発表した.また,大量のセグメンテーション画像を効率的に圧縮保存する手法の検討・開発を行った.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 金崎朝子	4. 巻 39
2. 論文標題 教師なし画像セグメンテーションのベーシックな手法と深層学習ベースの手法の紹介	5. 発行年 2021年
3. 雑誌名 MEDICAL IMAGING TECHNOLOGY	6. 最初と最後の頁 142-147
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kim Wonjik, Kanazaki Asako, Tanaka Masayuki	4. 巻 29
2. 論文標題 Unsupervised Learning of Image Segmentation Based on Differentiable Feature Clustering	5. 発行年 2020年
3. 雑誌名 IEEE Transactions on Image Processing	6. 最初と最後の頁 8055 ~ 8068
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TIP.2020.3011269	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 熊坂 雅之, 金崎 朝子
2. 発表標題 セグメンテーション画像を入力としたシーン分類手法の検討
3. 学会等名 Meeting on Image Recognition and Understandings (MIRU)
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------