

氏名（本籍）	むら さき かず ひこ 村 崎 和 彦（愛知県）
学位の種類	博士（工学）
学位記番号	甲第1005号
学位授与の日付	2019年3月19日
学位授与の要件	学位規則第4条第1項該当
学位論文題目	空間的配置情報を活用した教師なし物体領域抽出に関する研究

論文審査委員	（主査）教授 谷口 行信
	教授 赤倉 貴子 教授 古川 利博
	教授 八嶋 弘幸 准教授 寒水 孝司
	教授 浜本 隆之

論文内容の要旨

画像に写っている物体が何であるかを計算機によって認識させる画像認識の技術は近年急速に発展しており、人の目による認識精度を超えるほどの結果が報告されている。しかし、そうした高精度な画像認識器を学習するためには大量の正解ラベル付き画像を用意することが必須となっており、認識対象が増加するのに応じて認識器にとっての教師情報である正解ラベルを大量に用意しなければならないという問題が生じている。十分な量の正解ラベルを用意することは多大なコストがかかるが、一方で学習に用いることができる画像データ自体は全世界で爆発的に増えており、様々な物体が写る大量の画像を手に入れることのコストは日々低下していると言ってよいだろう。そこで、本論文ではこうした正解ラベルが与えられていない大量の画像を用いて新たな物体に関する情報を自動で獲得するような認識器の実現を目的とし、画像データのみから物体認識を行う教師なし物体領域抽出について述べる。

教師情報を用いずに画像を認識するにあたっては、多数用意された画像の中で同一の物体を示す画像領域がどこであるかを推測することが解決すべき課題となる。そのようにして同一物体を示す画像領域を集めることができれば、それらを教師情報として画像認識器を学習することができるだろう。この同一物体の領域を見つける手がかりとして、従来研

究では局所的な見た目の類似性が主に用いられていた。見た目の類似する領域が同一物体であると仮定することは有効なアプローチではあるが、見た目が類似していても異なる物体である場合や、同一物体であっても見た目が大きく異なる場合もあり、物体の形状変化や画像の撮影条件が大きく変わるような一般的な画像に対しては、適用が難しいという問題があった。これに対して本論文では、局所領域の類似性に加えて局所領域間の空間的な配置関係を手がかりとした領域抽出手法を提案する。空間的な配置関係とは、例えば左右に並んでいるといった画像上の配置関係だけでなく、前後に並んでおり一方が他方を遮蔽しているといった3次元の意味を含んだ配置関係を意味する。

本論文では空間的な配置関係を活用した教師なし物体領域抽出に関して、主に3つの提案を述べる。初めに、局所的な画像特徴の類似による手がかりと空間的配置関係に基づく手がかりを同時に考慮するための枠組みとして、生成モデルに基づく物体領域手法を提案する。更に、提案する生成モデルにおいてより効果的に見た目の類似性及び空間的配置関係を捉えるためのそれぞれの画像特徴の改良について提案する。

本論文は次のように構成される。第1章にて本論文の目的について述べる。第2章にて教師なし物体領域抽出における関連研究と課題について述べる。第3章では従来の画像特徴の類似に加えて空間的配置関係を加味した物体領域抽出のための生成モデルについて詳細を説明し、各種データセットにおける実験結果を示す。第4章では空間的配置関係をより効果的に用いるため、境界線における前後関係を明示的に考慮した特徴表現を提案し、前後関係の変化が生じる遮蔽境界線検出の実験によってその効果を示す。第5章では局所領域の類似性をより効果的に用いるため、大量の一般物体を識別するよう学習された深層特徴を用いて物体のカテゴリ分類に特化した特徴表現を提案し、単一の物体カテゴリを対象とした教師なし物体検出の実験においてその効果を示す。第6章では本論文の提案技術によって達成される教師なし物体領域抽出についてまとめ、その将来性について述べる。

第3章にて提案する教師なし物体領域抽出手法は、1枚の画像を **superpixel** と呼ばれる細かく分割された領域の集合として捉え、各 **superpixel** に対して何らかのトピックを付与するトピックモデルをベースとする。**Superpixel** を単位とするトピックモデルに、前述の空間的配置関係の手がかりを導入し、**superpixel** 間の境界に何らかの意味付けとして境界トピックを付与するモデルを提案する。提案モデルによって、特定のトピック間に現れる境界トピックの偏りと境界トピックから生成される境界画像特徴の偏りを捉えることができる。これは、例えば物体と背景との境界に輪郭らしい特徴が表れやすいといった傾向や人と自転車の領域は上下に並ぶように配置されやすいといった傾向を、教師情報を用いずに捉えるものである。単数物体及び複数物体を含むデータセットを用いた実験によって、従来のトピックモデルに比べて高精度に物体領域抽出を行えることを示す。

「領域間の類似性」及び「空間的な配置関係」が本研究において重要な手がかりとなるが、手がかりとして用いる画像特徴の抽出方法によってもその効果は異なる。従来は、SIFTのように撮影条件による見えの変化の影響を抑えた特徴表現を予め定められたルールによって抽出していたが、近年では何らかの学習用データを用いて学習によって適切な特徴表現を獲得するアプローチが多数取り組まれており、学習すべき対象を適切にデザインすることで有効な特徴表現が得られる場合が多い。本論文においても、物体領域抽出手法の精度向上を狙い、「領域間の類似性」を評価する特徴及び「空間的な配置関係」を示す特徴について学習に基づく新たな特徴表現を提案する。

第4章では、「空間的な配置関係」を示す画像特徴を得るため、遮蔽境界—境界線において領域が前後に隔絶されている境界—を検出し、遮蔽境界においてどちらが前方でどちらが後方であるかを推定する手法を提案する。遮蔽境界の検出を目的とした研究はこれまでも取り組まれているが、従来手法ではエッジの形状に基づく前後関係の推定とスペクトラルクラスタリングによる大域的な最適化が用いられており、局所的な推定の精度が低く、また大域最適化に時間がかかるという問題があった。これに対して提案手法では、局所的な領域での勾配情報や自己相関特徴などの画像特徴から境界線の形状と周辺の前後関係を一度に求めるアプローチを取る。境界線の形状推定と前後関係推定を機械学習によって効率的に認識するため、前後関係が付与された境界線形状を事前に一定数のクラスに分類し、画像特徴から各局所領域がどの形状クラスに属するかを識別する問題を設定し、その識別スコアによって各画素における境界線らしさとどちらが前景であるかの確率を算出する手法を提案する。このようにして得られる識別スコアは境界線の傾きだけでなく、周辺の前後関係を内包した特徴表現となる。画像特徴から境界線形状を識別する境界特徴表現を活用することで、局所的な遮蔽関係の推定精度が向上し、簡易な大域最適化処理によって高速かつ高精度に遮蔽境界検出が可能になる。多様な画像に対して遮蔽境界を検出する実験によって、提案する特徴表現を用いることで、従来手法と比較して遮蔽境界検出の精度が向上し、更に計算時間が削減することを示す。

第5章では、「領域間の類似性」を評価する画像特徴を得るため、大規模な物体識別問題を事前に学習させた深層特徴によって画像データセットから同一カテゴリの物体を見つけ出す手法を提案する。類似性を評価する画像特徴表現は、照明条件や撮影条件に対して影響を受けることなく、同様の物体に対して同様の特徴表現を獲得できることが望ましいが、そうした性質の特徴表現として、近年では深層特徴がよく用いられている。深層特徴は物体識別問題に対して事前に学習された多層ニューラルネットワークの中間出力を画像特徴として用いるアプローチであり、事前学習に用いるデータセットが十分に多様な認識対象・撮影条件を含んでいれば、事前学習の対象と異なる対象物体に対しても有効な特徴と

なり得ることが知られている。深層特徴を用いて教師情報を用いずに同一カテゴリの物体を抽出する手法も提案されており、様々なデータセットにおいて頑健に物体抽出ができることが示されている。従来手法では深層特徴の主成分分析に基づく物体の特徴抽出が用いられているが、このアプローチではデータセットに対して単一カテゴリの物体しか取り出すことができず、我々が目的とする多種の物体抽出には適用できない。多種の物体抽出のためには、適切な距離指標に基づいて深層特徴を多数の類似特徴にクラスタリングする必要がある。提案手法では、一般的な識別問題を学習した学習済み多層ニューラルネットの中間特徴には同一カテゴリを示す特徴ベクトルが放射状に連なる傾向があることに基づき、コサイン距離に基づく球面クラスタリングを導入する。畳み込みニューラルネットワークの中間特徴を入力画像の各画素に対応させ、それらを球面クラスタリングすることで得られるクラスタ番号を各画素の特徴表現として用いることができる。多数の画像から教師情報を用いずに同一カテゴリの物体を検出する実験によって、従来手法と比較して高精度に同一カテゴリの物体を抽出できることを示す。

論文審査の結果の要旨

学長からの審査の付託を受けて、標記6名の審査委員で構成する審査委員会を組織し、提出された学位論文について審査を行った。

審査委員会では、学位申請者より学位論文の内容、あるいは前回審査における指摘事項に対する対応結果について説明し、その後、質疑応答を実施することで、博士論文として満たすべき条件や必要な修正点を確認する、という形式で審査を行った。

本論文は、「空間的配置情報を活用した教師なし物体領域抽出に関する研究」と題し、全7章より構成されている。

画像認識技術は近年、急速に発展しており、人の目による認識精度を超えるほどの結果が報告されている。しかし、高精度な画像認識を実現するためには正解ラベル付きの教師データを大量に用意する必要があり、教師データの作成に多大なコストを要するという問題を抱えている。本論文では、教師なし物体領域抽出、つまり認識対象の物体に関する教師情報が付与されていない画像集合から物体領域を抽出する手法の開発を目的としている。

第1章では、物体領域抽出の代表的な問題設定について整理し、物体領域抽出の課題について述べている。

第2章では、本論文で扱う教師なし物体領域抽出の問題設定を定義し、前提条件を明らかにするとともに、提案する物体領域抽出のアプローチについて述べている。提案するアプローチは、まず各画像を **superpixel** とよばれる小領域に分割する、(1) 各

super pixel から見た目の類似性を評価するための領域特徴を抽出する, (2) 隣接する super pixel の空間的配置関係を評価するための境界特徴を抽出する, (3) 抽出した 2 種類の特徴を統合することで最終的な物体領域を抽出するものである. 特に, 境界特徴として, 形状だけでなく空間的配置関係 (上下・左右・前後) を考慮することを特徴としている.

第 3 章では, あらかじめ正解ラベルが付与された大量の画像データセット ImageNet を用いて学習した畳み込みニューラルネットワーク (CNN) から得られる中間出力, いわゆる深層特徴を球面クラスタリングすることにより領域特徴を抽出する手法を提案している. 画像集合から教師なしに同一物体位置を検出する co-localization とよばれる問題に対して提案法を適用し, 従来法を上回る精度を達成している.

第 4 章では, 領域間の空間的配置関係を捉えるため, 境界線における前後関係を明示的に事前学習させた境界特徴表現を提案している. 画像から物体輪郭となる遮蔽境界線を検出し, 物体の前後関係を推定する実験によってその精度を評価している.

第 5 章では, 確率的生成モデルをベースとして, 観測される画像特徴 (具体的には, 領域特徴と境界特徴) が潜在的な物体カテゴリから生成されることを表現したトピックモデルを提案し, 教師なし物体領域抽出問題に適用している. 各種データセットを用いた精度評価実験により提案モデルの有効性を示している.

第 6 章では, 第 3 章, 第 4 章で提案した領域特徴と境界特徴と, 第 5 章で提案したトピックモデルを統合することで, 物体領域抽出の精度が更に向上することを確認している. 入力として与えられる画像集合に複数カテゴリの物体が含まれる, より現実的な問題設定において, 対象物体の見た目が大きく変わるような難しいデータやノイズ混じりのデータに対しても物体領域抽出が可能であることを確認している.

第 7 章では, 以上の成果をまとめるとともに, 将来課題について述べている.

以上を要するに, 本論文の貢献は以下の三つである: 第一に, 教師なし物体領域抽出問題に対して, 画像領域の見た目の類似性, 及び, 隣接する領域間の空間的配置関係を捉えるために, 事前学習に基づく領域特徴表現と境界特徴表現をそれぞれ提案し, 実験により特徴表現の有効性を示していること, 第二に, 領域の類似性と空間的配置関係を同時に考慮するために, トピックモデルに基づく物体領域抽出法を提案し有効性を示したこと, 第三に, 単一カテゴリのみならず, 複数カテゴリの物体を抽出する複数カテゴリ物体領域抽出の問題設定で精度を評価し有効性を示したことである. これらの成果は, 爆発的に増加する画像・映像メディアを構造化する際に, 大幅なコスト削減をもたらす可能性を持つものであり, 情報工学・画像工学の分野において寄与するところが大きい.

以上により, 審査委員会は, 本論文を博士 (工学) の学位論文として十分に価値あるものと認める.