



Visual Learning Models for Real-world Object Recognition using Diverse Features

野宮, 浩揮

(Degree)

博士 (工学)

(Date of Degree)

2008-09-25

(Date of Publication)

2009-07-31

(Resource Type)

doctoral thesis

(Report Number)

甲4404

(URL)

<https://hdl.handle.net/20.500.14094/D1004404>

※ 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。



氏 名	野宮 浩揮
博士の専攻分野の名称	博士（工学）
学 位 記 番 号	博い第 537 号
学位授与の 要 件	学位規則第 5 条第 1 項該当
学位授与の 日 付	平成 20 年 9 月 25 日

【 学位論文題目 】

Visual Learning Models for Real-world Object Recognition using Diverse Features（実世界における物体認識のための多様な特徴を用いた視覚的学習法に関する研究）

審 査 委 員

主 査	教 授	上原 邦昭
	教 授	有木 康雄
	教 授	大川 剛直
	教 授	玉置 久

(氏名： 野宮 浩揮 NO. 1)

コンテンツベースの画像検索やカメラを利用した監視システムなど、近年の画像を用いたアプリケーションの増加に伴い、画像中の物体を正確かつ効率的に認識する手法の必要性が高まっている。特に、インターネットを介した画像検索など、大規模な画像データベースから様々な種類の画像を検索する場合には、実世界に存在する多種多様な物体を正しく識別する必要がある。しかしながら、現在の物体認識システムは、識別対象の物体を（例えば人の顔などに）限定すれば、高い認識性能を達成できるが、人間の視覚ように様々な物体を正確に識別することは非常に困難である。そこで、本論文では、コンピュータビジョンを用いた画像認識システムに機械学習の概念を導入し、より汎用的で高精度な物体認識手法を構築することのできる視覚的学習法を提案する。

本論文は六章で構成される。第一章では、視覚的学習法およびその基盤となる機械学習の観点から、本研究の背景と目的を示す。第二章では視覚的学習について述べる。従来のパターン認識等による認識手法における問題点と、問題解決のための、視覚的学習法による機械学習の導入について述べる。また、視覚的学習法を4種類のモデルに分類し、それぞれのモデルに対応する既存手法とその問題点、および提案手法との関連について言及する。

第三章では、視覚的学習法の基盤となる機械学習法の一つである Boosting 法について述べ、その中でも高いデータ分類性能を示すため広く用いられている AdaBoost アルゴリズムの分類性能の改善を行う。AdaBoost は、学習用に与えられたデータ（訓練データ）を用いて複数回学習を行い、各回で得られた複数の仮説を統合することによって、単一の仮説を用いる場合より高い分類性能を実現するアンサンブル学習と呼ばれる学習法の一つである。従来の AdaBoost は、誤ったクラスラベルが割り当てられている事例や、外れ値を持つ事例など、非常に分類が困難な事例（これらを **hard example** と呼ぶ）が訓練データに含まれている場合に、分類性能が極端に悪化することがあるという問題を抱えている。この問題の主な原因として、**hard example** をうまく分類しようとして必要以上に学習してしまい、結果として **hard example** の分類に特化した分類器が生成されてしまうことが挙げられる。この現象を過学習と呼ぶ。過学習を起こした分類器は、**hard example** は正しく分類するが、他の多くの事例を誤分類してしまい、全体的な分類性能が低下する。過学習は **hard example** に必要以上に高い重みが与えられることが原因で発生するので、過学習を防ぐために、AdaBoost の重み更新規則にしきい値を導入し、**hard example** の重みが高くなり過ぎないようにアルゴリズムを改良した NadaBoost アルゴリズムを提案する。また、NadaBoost では、学習の際に AdaBoost の理論的な誤差の上限を求め、NadaBoost の誤差上限が常に AdaBoost を下回るようにしきい値を設定することにより、AdaBoost を上回る分類精度を実現している。可能なしきい値の数は非常に多いため、ビームサーチを用いて適切なしきい値を効率的に決定している。ビームサーチは最良優先探索の一種であり、探索空間を制限し

(氏名： 野宮 浩揮 NO. 2)

て探索を行うことにより、通常の最良優先探索に比べて短時間で最適に近い解を得ることができる。機械学習のベンチマークデータセットを用いた分類実験では、**hard example** が多く含まれる場合だけでなく、**hard example** が比較的少ない場合であっても、提案手法は AdaBoost に比べて高い分類精度を実現している。

第四章および第五章では、視覚的学習法に基づく画像中の物体認識について述べる。視覚的学習法では、画像データから、物体を構成する画素の明度や物体の輪郭・テクスチャなどの特徴を抽出し、得られた特徴を用いて学習を行うことにより物体を認識する。まず第四章では、画像を構成している各画素の明度の特徴として用いて、Set Covering Machine (SCM) アルゴリズムに基づく物体認識法を提案する。画像中の各画素に対して SCM アルゴリズムを適用することにより、例えば「明度がしきい値以下であれば正事例である」といった分類規則が生成される。しかし、画像に含まれる画素数は非常に多いため、すべての画素について分類規則を生成すると計算量が極めて大きくなる。さらに、画像中には分類に寄与しない冗長な画素も多数存在するので、すべての画素を用いるより、分類に適している画素のみを用いる方が効率的に分類を行うことができる。そこで、提案手法では、各画素の明度の分布に対して判別分析を適用し、分類に寄与する度合いを求めることにより、多数の画素の中から少数の有用な画素を発見し、有用な画素のみを用いて分類規則を生成する。これにより、すべての画素を用いた場合と比べて、分類精度には影響を与えず、学習時間が著しく短縮される。また、訓練データとして与えられた画像データに平滑化や明瞭化などのフィルタリングを施し、新たな特徴を得ている。フィルタには様々な種類があるため、最適なフィルタの組み合わせを発見する必要がある。しかし、フィルタの組み合わせはフィルタの数に対して指数関数的に増大するので、ビームサーチと呼ばれる、探索空間を適切に制限することのできる探索法を用いて、最適（もしくは準最適）なフィルタの組み合わせを高速に探索している。これにより、学習にかかる時間を著しく増大させることなく、フィルタリングにより分類精度を向上できることが、画像の分類実験を通じて確認できている。

コンテンツベースの画像検索など、実用的なアプリケーションが対象とする画像データは、一般に実世界の複雑な物体を含んでおり、また背景や照明条件なども一定でないため、これらの画像データの認識において十分な精度を実現するためには、人の視覚システムが行っているように、より多くの特徴を適切に組み合わせる必要がある。そこで、第五章では、輪郭・色・テクスチャ・空間周波数など、多数の特徴を組み合わせた視覚的学習法を構成する。具体的には、一つまたは少数の特徴を用いて一つの識別器を生成し、アンサンブル学習法に基づいて複数の識別器を統合する。また、複数の識別器が互いに助け合いながら協調して学習を進めていく形で学習を行う。それぞれの識別器にとって正しい識別が困難な物体は、学習過程において、その識別器が誤分類した事例から判断することができる。ある識別器が頻繁に誤分類した事例は、その

(氏名： 野宮 浩揮 NO. 3)

識別器にとって識別困難であると考えられるので、その事例は他の識別器で重点的に学習する。その代わりに、その識別器は、他の識別器が誤分類した事例を学習する。このような学習ステップを繰り返すことにより、各々の学習器が相互に作用しながら相補的に学習することができる。この学習法により、それぞれの識別器は、自身にとって識別が比較的容易である事例の識別に特化されるため、単独で学習を行う場合よりも高い認識精度を実現できる。また、未知の事例を識別する際には、その事例を正しく識別できる可能性の高い識別器を推定し、その識別器の予測を重視することにより、信頼性の高い予測を得ている。どの識別器が正しい予測を与える可能性が高いかということは、学習過程において訓練事例に対して認識を行った結果得られる混同行列を用いて、どの種類の物体を正しく認識したか、あるいは誤って分類したかという傾向を知ることにより判断できる。また、この相補的学習法は、反復的な学習を行うため AdaBoost の学習機構に基づいて構成される。しかし、AdaBoost アルゴリズムには、過学習を起こしやすいという重大な問題点がある。そこで、過学習を抑制できる NadaBoost アルゴリズムを導入することにより、非常に複雑な背景を持つ画像や変形した物体・照明条件の悪い物体を含む hard example に対しても過学習を起こすことのない、頑健な認識手法を構成している。また、元来 NadaBoost は単一の学習アルゴリズムを用いたアンサンブル学習を行うように設計されているのに対し、相補的学習法では複数の学習アルゴリズムを用いて複数の識別器を同時に訓練するため、相補的学習に適用できるように NadaBoost アルゴリズムの拡張を行っている。これにより、実世界の画像データを用いた物体認識実験において、提案手法は単一あるいは少数の特徴を用いて物体認識を行う既存の手法と比べて高い認識精度を達成した。また、hard example を多く含むデータセットに対しても、過学習を起こすことなく、良好な認識性能が得られている。

最後に、第六章では、本研究をまとめ、今後の展望について述べる。本研究では、様々な特徴を用いて複数の識別器を構成し、それらを協調的に学習させることにより、様々な物体を識別できる手法を提案した。今後の展望としては、人の視覚的能力により近づくために、人の視覚システムにおける様々なモデルを導入することが挙げられる。

(別紙 1)

論文審査の結果の要旨

氏名	野宮 浩揮		
論文 題目	Visual Learning Models for Real-world Object Recognition using Diverse Features (実世界における物体認識のための多様な特徴を用いた視覚的学習法に関する研究)		
審査委員	区 分	職 名	氏 名
	主 査	教 授	上原 邦昭
	副 査	教 授	有木 康雄
	副 査	教 授	大川 剛直
	副 査	教 授	玉置 久
	副 査		
要 旨			
本論文では、画像や映像などのマルチメディアデータを用いた検索・認識システムを構成するために必要となる画像中の物体認識手法として、コンピュータビジョン処理に機械学習を導入した認識手法である視覚的学習法を提案している。また、視覚的学習に用いる機械学習法として、boosting アルゴリズムを改善した NadaBoost アルゴリズムを提案している。 本論文は六章で構成されている。第一章では、本研究の背景および目的を述べている。従来の物体認識システムは、顔や文字など限定された識別課題に対しては正確な認識を行うことが可能であるが、人間のように様々な物体を正確に識別することはできない。そこで本論文では、人の視覚的能力に匹敵する物体認識を行うことを目的として、機械学習法を導入することにより自動的に物体を識別する知識を獲得できる視覚的学習法を提案している。視覚的学習法により汎用的かつ正確な画像認識が可能となるため、視覚的学習に関する研究分野がさかんに行われている。しかし、現時点では人と比べて十分な認識性能を持つ視覚的学習法が存在しないため、研究の余地は未だ大いにあり、本研究は非常に有用であるといえる。 第二章では、まず視覚的学習法の概要について言及し、視覚的学習による、従来の（学習を用いない）認識手法における問題点の解決法について述べている。次に、視覚的学習法を4つのモデルに分類し、各モデルに該当する既存手法とその問題点について述べられている。また、提案手法のモデルの位置づけも示されている。 第三章では、本論文において提案されている視覚的学習法で用いられている機械学習法である AdaBoost 法の持つ問題点について述べ、問題解決のための手法を提案している。AdaBoost は、学習アルゴリズムを複数回実行し、各回で得られた仮説を統合してアンサンブル分類器を構成することにより、単体の仮説を用いた場合と比べて高精度での分類を行うことのできるアンサンブル学習法の一つである。従来の AdaBoost には、外れ値を持つ事例や、分類が非常に困難な事例（これらを hard example と呼ぶ）が訓練データに含まれていると、hard example は正しく分類できるが、それ以外の多くの事例を誤分類してしまう仮説が生成される、過学習と呼ばれる現象を引き起こすという問題点がある。過学習が引き起こされると、一般に分類性能が著しく低下する。この問題を解決するために、AdaBoost アルゴリズムの重み更新規則を改善し、NadaBoost アルゴリズムを提案している。過学習は hard example に過剰な重みを与えてしまうことが原因で発生するので、NadaBoost アルゴリズムでは重みに閾値を設け、重みが閾値を超えた事例を hard example とみなし、誤分類されても重みを減らす。可能な閾値の数は非常に多いので、ビームサーチを導入し、最適に近い閾値を低い計算コストで発見できるようにしている。機械学習手法の検証に広く用いられるベンチマークデータセットを用いた実験では、hard example が多く含まれる場合だけでなく、hard example が比較的少ない場合でも、AdaBoost に比べて高い分類精度を実現できることが示されている。 第四章および第五章は、それぞれ視覚的学習法が提案されている。まず第四章では、画像を構成する個々の画素の明度の特徴として用いた視覚的学習法が提案されている。学習用に与えられた画像データにおける画素の明度分布に対し Set Covering Machine アルゴリズムを適用することにより、未知の画像を識別するための識別規則を自動的に生成している。画像には多数の冗長な画素が含まれているため、すべての画素を用いると計算コストが増大し、認識精度も低下するので、少数の有用な画素を発見するための特徴選択法を提案し、効率的な認識を行っている。各画素の明度分布に対して級間分散と級内分散の比をその			

氏名	野宮 浩揮
画素の有用度と定め、有用度の高い少数の画素を用いることにより、識別精度を低下させることなく、計算コストを大幅に減少している。また、学習用画像に対して平滑化や明瞭化、輪郭抽出などのフィルタを適用し、新たな特徴を得ている。フィルタの組み合わせは非常に多数であるため、ブームサーチを用いて最適なフィルタの組み合わせを効率的に探索している。これにより、計算コストを抑えて正確な識別を行えることが、画像の分類実験を通じて示されている。 上述の視覚的学習法では画素の明度のみを特徴として用いているため、実世界の複雑な物体の認識に対しては認識精度が不十分である。そこで、第五章では、アンサンブル学習法に基づき、物体の輪郭やテクスチャ、空間周波数など、多数の特徴を組み合わせることで認識に用いることのできる視覚的学習法を提案している。一つの特徴に対して一つの識別器を生成し、それらを統合してより識別性能の高いアンサンブル識別器を構成する。また、提案手法では、各識別器を個別に訓練する従来のアンサンブル学習法と異なり、識別器間の相互作用を導入している。具体的には、学習過程において、各識別器は自身にとって正しい識別が困難な事例を他の識別器に学習してもらう代わりに、他の識別器にとって識別困難な事例を学習する。このような相補的な学習法により、従来のアンサンブル学習法よりも高い認識精度を実現している。提案手法は AdaBoost アルゴリズムに基づいて学習を行うが、AdaBoost は hard example（変形した物体や照明条件の悪い物体を含む画像、非常に複雑な背景を持つ画像などに相当する）により過学習を引き起こすという問題点を持つため、第三章で提案した NadaBoost アルゴリズムを視覚的学習に適用できるよう拡張し、過学習を防止している。これにより、hard example を多く含む実世界の画像データを用いた物体認識実験において、提案手法は既存の認識手法と比べて高い認識精度を達成できることが示されている。 第六章では、本研究をまとめ、今後の展望について述べられている。本研究は複雑化・多様化する近年の画像データを用いた様々なアプリケーションに対して適用可能な物体認識手法について研究したものであり、機械学習を導入することによって、従来の認識手法に比べて高い汎用性・認識精度を実現している点から、物体認識手法における視覚的学習法の有効性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、学位申請者の野宮浩揮は、博士（工学）の学位を得る資格があると認める。	