

## [ショートペーパー]

# 物体属性に関する論理制約を用いた車載カメラ画像からの物体検出

星谷 那月<sup>†</sup> 出口 大輔<sup>†</sup> 陳 嘉雷<sup>†</sup> 村瀬 洋<sup>†</sup> 伊藤 誠悟<sup>†</sup>

<sup>†</sup>名古屋大学

**あらまし** 本研究では、車載カメラ画像を対象とした物体検出器の精度改善を目的とし、論理制約を活用した半教師あり学習手法を提案する。一般的な半教師あり学習では、アノテーションがないデータに対するモデルの出力を正解データとして用いる。しかし、モデルの出力をそのまま用いた場合、誤検出も含めて正解データとしてアノテーションしてしまうという課題があった。提案手法では、論理制約を満たす推論結果を選択的に学習データとして使用することでデータセットの質を改善する。これにより、物体検出器の精度改善を図る。ROAD-R データセットを用いた実験では、平均適合率の平均 (mAP) はベースライン手法を上回り、論理制約の充足度 (RSR) の改善は限定的であった。

**キーワード** 物体認識, 物体検出, 自動運転, 半教師あり学習, 論理制約

## Object Detection using In-Vehicle Camera Images with Object Attributes Constraints

Natsu HOSHIYA<sup>†</sup>, Daisuke DEGUCHI<sup>†</sup>, Jialei CHEN<sup>†</sup>, Hiroshi MURASE<sup>†</sup>, and Seigo ITO<sup>†</sup>

<sup>†</sup> Nagoya University

### 1. まえがき

近年、自動運転技術の発展により物体認識の精度は向上してきている。しかし、物体の状態や位置を認識する際に現実的な制約を無視することがあるという問題がある。例えば信号の状態を「赤であり青である」と認識した場合、自動運転車は正常に動作しない。このような現実的な制約を無視した誤認識は重大な事故につながるおそれがある。さらに、物体認識モデルの学習には膨大なデータと高コストなアノテーション作業が必要であり、多様な環境下でのデータ収集には限界がある。そこで本研究では、物体属性に関する論理制約を活用し、アノテーションがないデータも有効に活用できる半教師あり学習の手法を提案する。これにより、少数のアノテーション付きデータと論理制約を活用し、高精度な物体認識モデルの構築を目指す。

本報告では、ROAD-R データセット [1] を用いる。これは、物体の情報に加え、交通環境における現実的な制約が論理制約として与えられた自動運転用の初となるデータセットである。ROAD-R データセットには、物体の種類、物体の状態、物体の位置という3つの情報が手動でラベル付けされている。これらに対応するラベルは全部で41種類あり、さらにこれらのラベルに基づいた243の現実的な論理制約が設定されている。

### 2. 提案手法

本報告では、半教師あり学習の手法として、アノテーションが付与されていないデータに対するモデルの出力を学習時の正解ラベルとして活用する。しかし、アノテーションのないデータを一旦推論し、その推論結果すべてを学習に用いると誤ったラベルが多く含まれる可能性があるため、論理制約を活用して信頼できるデータを選別する手法を提案する。

まず、ROAD-R データセットのアノテーション付きデータを学習用・検証用・評価用に分割し、学習用データを「物体の状態」と「物体の位置」の組み合わせごとに分類し、それらの数が多い順にソートする。これを「ソート済み学習データ」として扱う。なお、アノテーション付きデータはすべて論理制約を満たしている。

最初に、アノテーション付きデータのみを用いて物体認識モデルを学習し、そのモデルをベースラインモデルとする。次に、このベースラインモデルをアノテーションなしデータに適用し、ソート済み学習データの1番目（最も多く現れる論理制約の組み合わせ）に該当する推論結果のみを学習データに追加して再学習を行う（1回目の学習）。

その後、 $n$  回目の学習では、 $n-1$  回目の学習モデルをアノテーションなしデータに適用し、ソート済み学習データの1番



図1 再学習後のモデルの検出結果の例

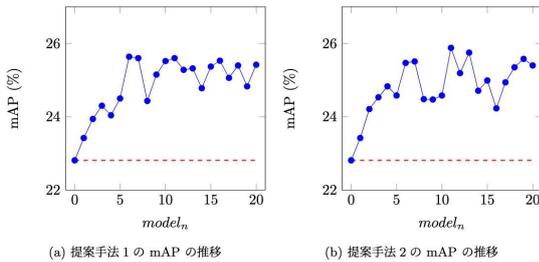


図2 手法1と2のmAP

目から  $n$  番目に該当する推論結果を学習データに追加する、この手法により、徐々に推論結果を拡張しながら学習を進める。

また、データの追加方法として、過去に追加したデータを新たな推論結果で置換する手法1と、過去のデータを保持しつつ新たな推論結果を追加する手法2を比較する。

### 3. 実験および考察

本報告では、既存の物体検出手法である DINO [2] により物体検出を行った後、DINO の Object Query を入力とする MLP によって詳細に物体認識を行う手法を採用した。

まず DINO で ROAD-R の全画像から 400 個のバウンディングボックスと特徴量を抽出し、正解バウンディングボックスと DINO の出力を IoU に基づき対応付ける。そして、最も IoU が大きいものを MLP の入力データとするようなデータセットを作成した。このデータセットを用いて学習を行い、2. で述べた方法で追加データを選別して繰り返し学習を行った。再学習後のモデルの出力の例を図1に示す。

評価指標には mAP (Mean Average Precision) と独自に定義した RSR (Requirements Satisfaction Rate) を用いた。ここで mAP は、物体の種類・状態・位置の 41 個のラベルの AP の平均値とした。RSR は推論結果が論理制約をどの程度満たしているかを示す指標であり、243 のルールをすべて満たす推論結果の割合として計算した以下の式で表される。

$$RSR = \frac{\text{制約を満たす推論結果の数}}{\text{すべての推論結果の数}}$$

mAP の結果を図2に、RSR の結果を図3に示す。

図2は、どちらの提案手法も mAP はベースラインモデルを常に上回っていることを示す。赤色の点線は、ベースラインモ

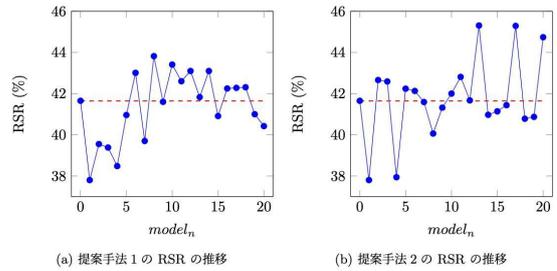


図3 手法1と2のRSR

デルの値である。また、図3は、手法1は、8回目の学習以降でベースラインを上回ることが多いのに対し、手法2では学習回数による変動幅が大きく、ベースラインを上回ることもあれば下回ることもあることが分かる。

提案手法が mAP でベースラインモデルを上回った要因は、「物体の状態」と「物体の位置」の組み合わせごとに分類したソート済み学習データをもとに推論データを選定したことにある。この手法により、物体の状態と物体の位置の AP が向上し、それに伴い物体の種類、物体の状態、そして物体の位置の AP をすべて平均した全体の mAP も上昇したと考えられる。

また、手法1において RSR が8回目の学習以降でベースラインモデルを上回ることが多かった要因として、学習を繰り返すことでモデルの認識精度が向上し、推論データの選定基準がより適切に機能するようになったことが考えられる。一方、手法2において RSR の変動幅が大きかった要因として、手法2では学習を進めるにつれて追加できる推論データが少なくなり、一回の学習で追加されるデータが限定されることが上げられる。制約ごとに推論での満たされやすさが異なるため、RSR が高くなった場合は、比較的満たされにくい制約を多く含むデータが追加された可能性がある。

### 4. むすび

本報告では、車載カメラ画像を対象とした物体検出器の精度向上を目的として、論理制約を活用した半教師あり学習手法を提案した。実験の結果、提案手法は mAP においてベースラインを常に上回る精度を示し、RSR についても8回目の学習以降でベースラインを上回る傾向が確認された。これにより、本手法が物体認識精度を向上させる有効なアプローチであることが示された。

今後の課題として、RSR のさらなる向上や、より広範なデータセットを用いた検証、リアルタイムシステムへの適用を検討する。

謝辞 本研究の一部は JSPS 科研費 23H03474 による。

### 文 献

- [1] E. Giunchiglia, S. Khan, F. Cuzzolin and T. Lukasiewicz, "ROAD-R: The Autonomous, Driving Dataset for Learning with Requirements", Machine Learning Journal (2023).
- [2] H. Zhang, F. Li, S. Liu, L. Zhang, H. Su, J. Zhu, L. Ni, and H. Shum, "DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection", In The Eleventh International Conference on Learning Representations, 2022b.