

ロボット工学とオートメーション分野で最高峰の国際会議 ICRA2024 にて路上

障害物検出に関する論文が採択

トヨタ自動車株式会社
社会システム P F 開発部
InfoTech-AS データ解析 1 グループ
野口千尋
大串俊明
山中正雄

トヨタ自動車株式会社は、画像から路上障害物を検出する AI 技術を開発しました。ロボット工学とオートメーション分野で最高峰の国際会議である ICRA2024 (International Conference on Robotics and Automation 2024) にて、当社の研究論文 "Road Obstacle Detection based on Unknown Objectness Scores" が採択され、2024 年 5 月 14~16 日の本会議にて発表を行います。

背景

安全・安心・快適な走行環境の実現に向けて、道路上の走行の妨げとなる障害物を自動で検出する技術は非常に重要です。同技術は、自動運転の場合において欠くことができないものですが、人が運転する場合においても、障害物情報の他車との共有、道路管理会社への情報提供など、様々なサービスの実現が可能になります。そこで、我々は、ドライブ・レコーダーなどの車載カメラ画像のみから路上障害物を検出することを主眼としました。

画像から路上障害物を検出する際の一番の課題は、障害物となりうる対象が多種多様な点です。道路上には車両からの様々な落下物、動物の死骸、落石、ポットホールなど、多岐に渡る障害物が出現するため、これらを網羅するように検出モデルを学習することは容易ではありません。加えて、道路上に障害物が落下している状況に遭遇する頻度は極めて低く、十分な学習データを収集することも困難です。そのため、提案法を含む多くのアプローチでは、異常検出手法を導入し、正常な走行環境においては起こりえない異常な事象を検出することで、路上障害物の検出に繋げるアプローチを採用しています。

路上障害物を検出するための最も素直なアプローチは物体検出手法を適用することです。物体検出手法では、あらかじめ物体ラベルを定義し、画像中の物体にラベル付け (アノテーション) することで、学習データセットを作成します。しかし、路上障害物を含む未知の物体はアノテーションされないため、これらは物体ではなく背景として学習されます。この結

果、画像中の路上障害物は、背景として学習され、学習して得られた検出モデルを用いて障害物を検出することが出来ません。このような背景から、多くの既存手法は、物体検出手法ではなく、意味的領域分割 (semantic segmentation) 手法をベースに構築されています。車載カメラで得られた画像中のほぼ全てのピクセルに対し、事前に定義したラベルを割り当てることのできるため、背景領域には、適切な背景ラベル (空, 道路など) が割り当てられ、事前にアノテーションされない物体領域が背景として学習されることはありません。また、車載カメラ画像の構成要素は種類に限られるため、比較的少量のラベルだけを用いて、画像中のほとんどの領域をカバーすることが可能です。そのため、事前に定義されたラベルのどれにも当てはまらない領域を、異常を含む領域として見なすことが可能となります。

しかし、上述の様な検出対象のラベルデータを明示的に利用しない教師なし学習では、特に車載カメラ映像の様な複雑な構造が含まれる画像において、安定的な性能を発揮することが難しく実用性の観点で課題があります。そこで我々は、教師あり学習可能な物体検出手法の考え方を異常検出手法に導入することで、より実用的な検出性能を達成可能な障害物検出手法を開発しました。

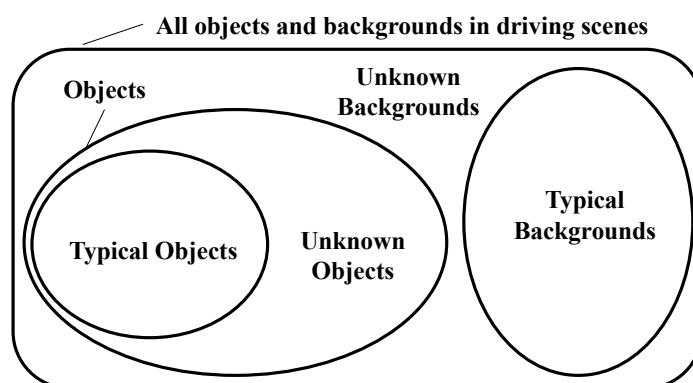


図 1 未知物体スコア (式 (1)) を表すベン図

技術概要

多くの物体検出手法では、物体スコアと呼ばれる物体らしさを表す尺度が利用されます。深層学習ベースの手法では、物体スコアも学習対象となり、より広い範囲の物体に汎化した物体らしさの獲得を目指します。路上障害物検出における検出対象は、道路上の物体に限られるため、物体スコアを上手く活用することで検出性能を向上させられると考えられます。図 1 に提案手法の基本的な考え方を表すベン図を示します。前述の様に、適用対象を車載カメラ映像に限定する場合、画像中に出現しうる要素は限られますが、これらの要素は大きく、物体 (車両, 人, ベンチなど) と背景 (道路, 空, 歩道など) の 2 つに分けられます。加えて、事前に定義したラベルに該当する要素を典型的な要素, 該当しない要素を未知の要

素と定義することで、車両や人などのラベルに含まれる物体は典型的な物体、ラベルに含まれない路上障害物は未知の物体と見なすことが出来ます。典型的な要素は、定義されたラベルを用いた教師あり学習を行う事ができ、物体スコアも物体検出の方法論に従うことで学習可能です。従って、未知の物体は、「物体だが、典型的ではない物体」(図1の青色の領域)として検出することが出来ます。具体的には、以下の様に未知物体スコアを定義します。

$$S_i = p_i^0 \prod_{k=1}^K (1 - p_{ik}) \quad (1)$$

ここで、 S_i はピクセル i の未知物体スコアを示します。 p_i^0 と p_{ik} はそれぞれピクセル i に対応する物体スコアとラベル k の予測確率を表します。 K は事前に定義したラベルの数です。未知物体スコアが高い領域を未知物体が含まれる領域として検出します。

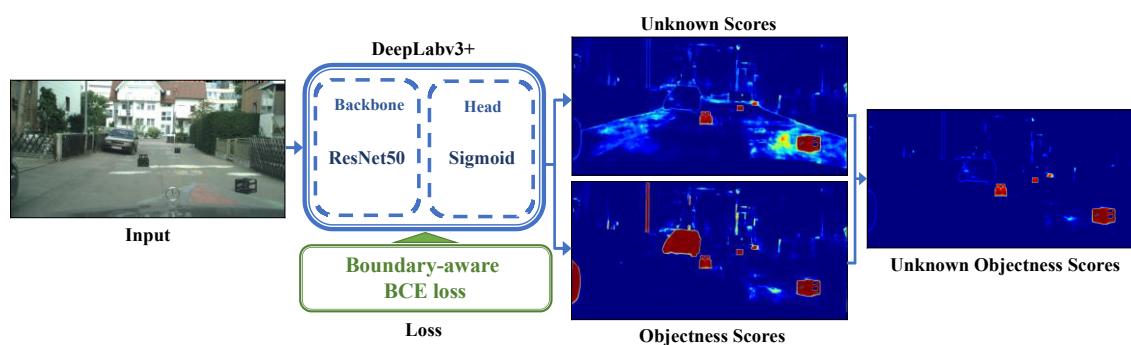


図 2 提案方法の概略図

図2に、提案手法の概略図を示します。与えられた画像は、標準的な意味的領域分割のためのニューラルネットワークに入力されます。提案法では、一部、標準的なネットワークとは異なる構成を採用しており、ヘッド部分にシグモイド関数を用いています。シグモイドヘッドは多ラベル分類を行う際に用いられる構成で、事前に定義したラベルのどれにも該当しない状態を自然に表現することが出来ます。同様の事を標準的に採用されるソフトマックス関数で実行する場合、未知の要素に該当するピクセルの予測確率のエントロピーを大きくする様に学習を行う必要がありますが、これは意味的領域分割の性能を大きく低下させてしまいます。加えて、シグモイドヘッドを採用することで、事前定義のラベルと物体スコアを同時に学習する事が出来るという利点もあります。

図2右では、物体スコアと未知スコアを可視化した結果を示しています。未知スコアは、式(1)において、物体スコアを除いた因子を表します。未知スコアだけでは、主に背景領域において多くの偽陽性が確認されますが、これと物体スコアを乗算した未知物体スコアでは、これらの大部分が抑制されていることが分かります。

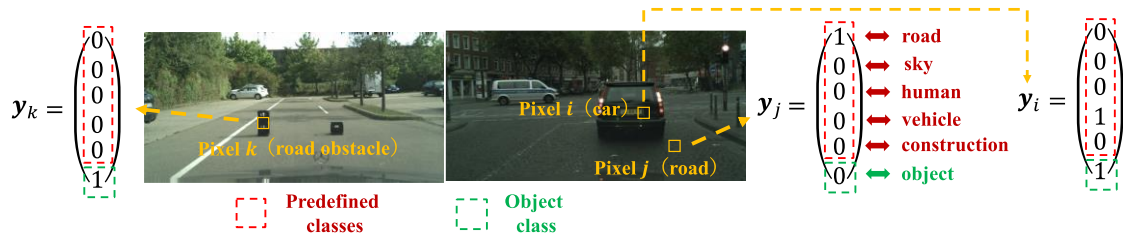


図 3 提案法におけるラベルの割り当て方

冒頭で述べた様に、通常の走行の中で路上障害物に遭遇する頻度は極めて低く、学習データとして十分な数のデータセットを収集することは困難です。しかし、少量であれば手に入れることは可能で、これら少量の学習データを活用して検出性能の改善を行うことが出来るということは実用上重要です。提案法では、物体スコアの検出性能向上のためにこれら少量の学習データを利用することが出来ます。図 3 左の画像中に、2 つの路上障害物があります。これらのピクセルに割り振られるラベルは、図中の y_k の様に、事前定義のラベルは全て 0、物体スコアに対応するラベルを 1 とします。これにより、新しい障害物が発見される度に新しいクラスを作成する必要なく、少量の学習データを有効に活用して検出性能の改善に繋げることが出来ます。

$$\mathcal{L}_n = -\frac{1}{N} \sum_{i=1}^N \sum_{k \in \mathcal{C}} f(y_{ik}, p_{ik}) - \frac{\lambda}{\sum_i \delta_i} \sum_{i=1}^N \delta_i \sum_{k \in \mathcal{C}} f(y_{ik}, p_{ik}) \quad (2)$$

式 (2) に、ネットワークを最適化するための目的関数を示します。 $\sum_{k \in \mathcal{C}} f(y_{ik}, p_{ik})$ は、ピクセル i におけるラベル y_{ik} と予測確率 p_{ik} の間のバイナリクロスエントロピーを示します。また、 \mathcal{C} は事前定義のラベルと物体クラスを合わせたラベル集合を示します。しかし、通常のバイナリクロスエントロピーによる最適化では、異なるラベル間の境界領域において多くの偽陽性が発生します。境界領域においてネットワークは、少なくとも 2 つのラベルの間で判断を迷うことになり、その結果境界領域における未知スコアが増加してしまうことが原因です。そこで、提案方法では、境界領域に対し追加のペナルティ (右辺第 2 項) を導入し、境界領域においても低いエントロピーの予測確率が出力される様に学習を行います。

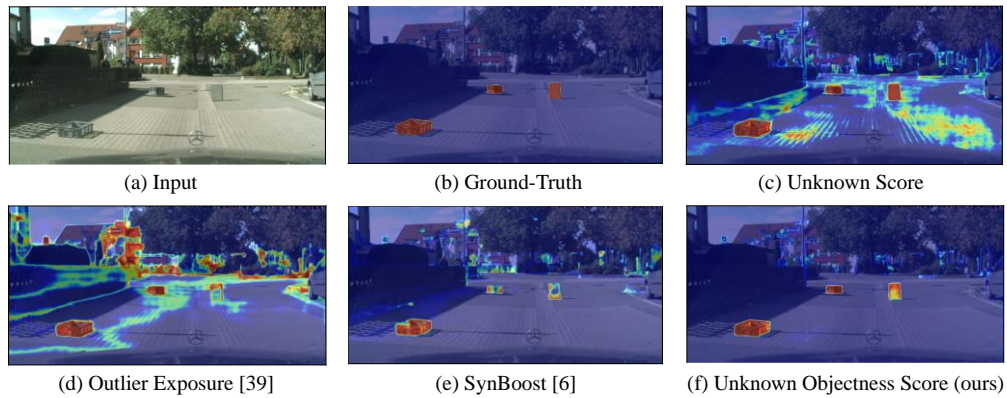


図 4 従来手法との定性的な比較

図 4 には、従来手法との比較を行なった定性評価の結果を示します。画像の中央付近に設置されている 3 つのオブジェクトを、提案法 (f) では比較的上手く検出出来ていることが分かります。従来手法と比較すると、主に背景領域において大幅に偽陽性の検出が抑制されていることが確認できます。

まとめと今後の課題

本研究では、物体検出と異常検出の 2 つの考え方を取り入れた新しい路上障害物検出手法を提案しました。教師あり学習が可能な物体スコアを活用することで大幅に検出性能を向上させることに成功し、特に背景領域における偽陽性の検出を大幅に抑制することが出来ました。一方で、物体スコアを導入したことは、提案モデルが障害物に対し高い物体スコアを予測することに失敗すると検出性能の低下に繋がることを意味します。そのため、より汎化性の高い物体スコアを学習することは更なる検出性能の向上の鍵になると考えられます。