

1. はじめに

非剛体で形状が多様な物体クラスの検出に対して、セマンティックセグメンテーションは有効な手段の一つである。しかしながら、正解ラベルを画素単位で付与するためアノテーションコストが高い。そこで本研究では、物体検出を併用して、未ラベルデータに対してアノテーションを自動的に行うフレームワークを構築する。そして、疑似的にアノテーションしたデータを用いて学習することで、不完全なデータ状況下における落雷検出モデルの高精度化を目指す。

2. 部分的教師ありマルチタスク学習

マルチタスクにおいて、全タスクの正解ラベルが揃わないデータでは、損失計算ができず学習が困難となる。この課題に対し、欠損ラベルを補完する部分的教師あり学習が提案されており、代表的な手法に BoMBo[1] がある。BoMBo の構造を図 1 に示す。BoMBo は、Backbone と Neck の Encoder と各タスクの Head で構成される。学習時、バウンディングボックスのみのデータには Box-for-Mask を行う。ここでは、GrabCut による疑似マスクを教師信号とするほか、Box 型マスクとアテンションマップの MSE を最小化して Segmentation head を学習する。一方、マスクのみのデータには Mask-for-Box を適用する。ここでは、マスクから外接矩形を算出して疑似ラベルとし、Detection head を学習する。

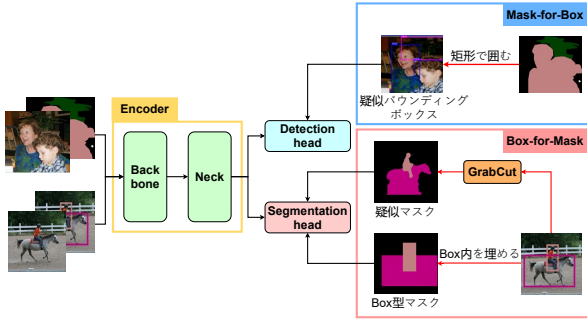


図 1: BoMBo の構造

3. 提案手法

本研究では、未ラベル画像とマスクのみの画像に対して両タスクの疑似ラベルを自動生成するフレームワークを構築する。そして、本フレームワークにより生成された疑似ラベルを用いて BoMBo による落雷検出の高精度化を図る。

3.1. 疑似ラベル生成フレームワーク

提案する疑似ラベル生成フレームワークを図 2 に示す。まず、未ラベル画像に対し、落雷画像でファインチューニングを行った YOLOv8 を用いて検出バウンディングボックスを生成する。次に、Segment Anything in High Quality (HQ-SAM)[2] を用いて疑似マスクを生成する。この際、事前学習済み重みを固定し、LoRA により追加された少数のパラメータのみを学習させることで、計算コストを抑えつつモデルを落雷画像に適応させる。

3.2. 部分的教師あり学習の適用

部分的教師あり学習の手法である BoMBo をもとに、検出バウンディングボックスから Box 型マスクを生成する。マスクのみのデータに対しては Mask-for-Box により疑似バウンディングボックスを補完し、これらを用いたマルチタスク学習を実施する。

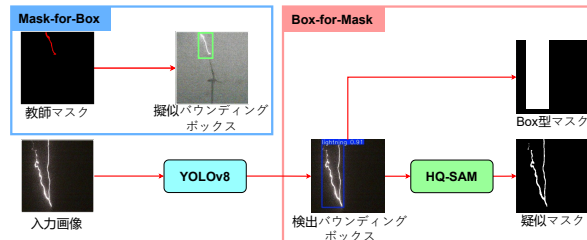


図 2: 疑似ラベル生成フレームワーク

4. 評価実験

画像処理や既存の基盤モデルを用いた疑似マスク作成と比較し、構築したアノテーションフレームワークによって生成されたラベルの有効性を検証する。

4.1. 実験概要

本実験では、学習用データとして教師マスクありデータ 56 枚、検出バウンディングボックスおよび疑似マスクありデータ 9,660 枚を用いる。評価用データは両タスク共通の 40 枚をデータセットとして用いる。モデルは RetinaNet、バッチサイズは 32、学習回数は 80epoch、評価指標は物体検出は AP、セグメンテーションは IoU とする。

4.2. 実験結果

定量的評価を表 1 に示す。表 1 より、LoRA を適用した HQ-SAM が最も高精度であり、通常の HQ-SAM と比較して、雷道の IoU が 8.75pt 向上した。

表 1: 定量的評価 [%]

LoRA	Method	AP	IoU	
			lightning	background
-	二値化	60.08	34.07	98.31
-	GrabCut	58.43	42.86	97.96
-	SAM	62.18	61.98	99.02
-	HQ-SAM	60.15	63.31	99.10
✓	HQ-SAM	65.99	72.06	99.41

HQ-SAM および LoRA を適用した HQ-SAM により生成したマスク画像を学習に使用した際のセグメンテーションの定量的評価を図 3 に示す。図 3 から、LoRA を適用することで雷道の詳細を捉えることができた。図 4 に物体検出結果を示す。図 4 より雷道の検出が可能であることを確認した。

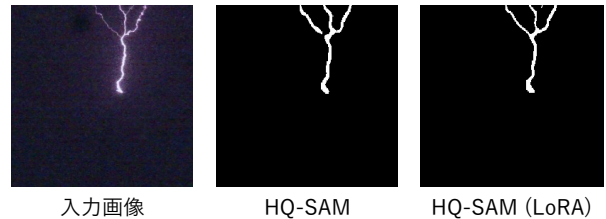


図 3: セグメンテーションの定量的評価

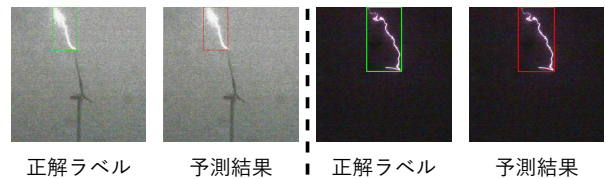


図 4: 物体検出の定量的評価

5. おわりに

本研究では、未ラベルデータを有効活用するため、自動アノテーションフレームワークを構築し、疑似ラベルを用いた部分的教師あり学習による落雷検出の高精度化を実現した。実験結果より、LoRA を適用した HQ-SAM による疑似マスクの導入が、落雷検出およびセグメンテーション精度の向上に有効であることを確認した。今後は、より高精度な疑似ラベル生成に向けた手法の検討を行う。

参考文献

- [1] H.Lê *et al.*, “Box for Mask and Mask for Box: weak losses for multi-task partially supervised learning”, BMVC, 2024.
- [2] L.Ke *et al.*, “Segment Anything in High Quality”, NeurIPS, 2023.