

## 1. はじめに

自動運転車が安全に走行するには、自車周辺の物体の 3 次元位置情報が必要である。その代表的な手法として 3D 物体検出が用いられるが、高性能化に伴いモデルの大規模化が課題となっている。この課題を解決する手法として、モデルをコンパクト化する枝刈りがある。枝刈りは構造化枝刈りと非構造化枝刈りの 2 つに大別される。非構造化枝刈りは重み行列を要素単位で削除することで、ネットワーク構造を維持したまま大幅にパラメータを削減できるという利点がある。この利点を踏まえ、本研究では非構造化枝刈りの物体検出モデルへの有効性を検証するとともに、枝刈りによる精度低下を抑制するために特定クラスおよび距離の重み付けを導入した非構造化枝刈りを提案する。

## 2. 先行研究

非構造化枝刈りの評価値として SNIP が用いられる。SNIP は損失関数に影響を与える重みを削除する手法であり、その評価値は、損失関数を  $\mathcal{L}$ 、 $i$  番目の重みを  $w_i$  として  $S_{\text{SNIP}}(w_i) = \left| \frac{\partial \mathcal{L}}{\partial w_i} w_i \right|$  のように示す。SNIP による枝刈りでは特定クラスの検出精度が低下することがある。先行研究 [1] では、特定クラスの精度低下を抑制するため、特定クラスを含む入力  $x^{\text{specific}}$  に対する損失から  $S_{\text{class}}(w_i) = \left| \frac{\partial \mathcal{L}(x^{\text{specific}})}{\partial w_i} w_i \right|$  のように示す。そして、 $S_{\text{SNIP}}$  と  $S_{\text{class}}$  を統合し枝刈りを行う。最終評価値  $S(w_i)$  を式 (1) に示す。

$$S(w_i) = S_{\text{class}}(w_i) + S_{\text{SNIP}}(w_i) \quad (1)$$

## 3. 提案手法

自動運転の観点では近距離領域での検出失敗が交通事故に直結するため、特定クラスに加え、近距離での検出精度低下を抑制する必要がある。そこで本研究では代表的な 3D 物体検出モデルである BEVFormer[2] を対象に、精度低下を抑制する特定クラスおよび距離の重み付けを導入した非構造化枝刈りを提案する。提案手法の概要を図 1 に示す。

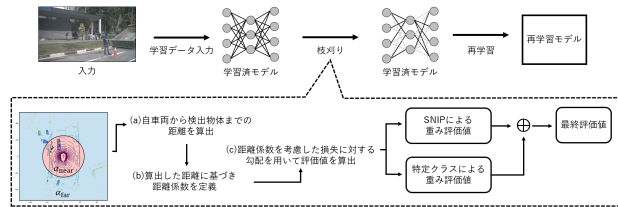


図 1: 提案手法の概要

提案手法では、距離係数  $\alpha(d)$  を付与した損失関数  $\mathcal{L}_{\text{dist}}$  を用いる。距離係数  $\alpha(d)$  を式 (2) に示す。

$$\alpha(d) = \alpha_{\text{far}} + (\alpha_{\text{near}} - \alpha_{\text{far}}) \exp(-d/\tau) \quad (2)$$

ここで、 $d$  は自車両から検出物体までの距離である。また、 $\alpha_{\text{near}}$ 、 $\alpha_{\text{far}}$  は近距離、遠距離に対する重みであり、 $\tau$  は距離に対する減衰の強さを表す。さらに距離に対する重み付き損失関数  $\mathcal{L}_{\text{dist}}$  を式 (3) に示す。

$$\mathcal{L}_{\text{dist}} = \frac{1}{N} \sum_{j=1}^N \alpha(d_j) \mathcal{L}_j \quad (3)$$

ここで、 $\mathcal{L}_j$  はサンプル  $j$  の検出損失、 $d_j$  はサンプル  $j$  までの距離、 $N$  はサンプル数を表す。式 (3) により、近距離ほど距離係数が大きくなり、損失への寄与が増加する。この距離に対する重み付き損失に基づき枝刈りを行い、評価値  $S(w_i)$  を式 (4) に示す。

$$S(w_i) = S_{\text{class}}^{\text{dist}}(w_i) + S_{\text{SNIP}}^{\text{dist}}(w_i) \quad (4)$$

ここで、 $S_{\text{class}}^{\text{dist}} = \left| \frac{\partial \mathcal{L}_{\text{dist}}(x^{\text{specific}})}{\partial w_i} w_i \right|$  は距離重み付き特定クラス評価値、 $S_{\text{SNIP}}^{\text{dist}} = \left| \frac{\partial \mathcal{L}_{\text{dist}}}{\partial w_i} w_i \right|$  は距離重み付き SNIP 評価値である。式 (4) により、近距離物体の検出に寄与する重みが削除されず、近距離性能の劣化抑制が期待される。

## 4. 評価実験

BEVFormer に対して提案手法の有効性を検証する。定量評価では、枝刈り前後の精度を全距離/近距離 (0~20 m) で比較する。定性評価では、検出結果の変化を確認する。

### 4.1. 実験条件

本実験では Backbone に ResNet101-DCN を用いる。学習設定はエポック数 24、バッチサイズ 64、枝刈り率 50%、70%、最適化は AdamW (lr=2e-4, weight decay=0.01) とし、データセットは NuScenes dataset を用いる。

### 4.2. 定量評価

全距離における枝刈りなしとありの精度を表 1、近距離における枝刈りなしとありの精度を表 2 に示す。表中の Methods は枝刈り時に距離係数を適用する距離帯 (全距離/近距離) を表す。表 1 より、枝刈り率 70% では SNIP は重要な重みまで削除されるが、SNIP/ours(全距離) は特定クラスの重みを保持するため、SNIP および SNIP/ours(近距離) を上回った。また、表 2 より、枝刈り率 50% の近距離評価では、SNIP/ours(近距離) が近距離検出に重要な重みを保持するため、SNIP や SNIP/ours(全距離) を上回った。

表 1: 全距離における定量評価

枝刈り率	Methods	NDS	mAP	mATE	mASE	mAOE	mAVE	mAAE
0%	枝刈りなし	0.524	0.417	0.655	0.273	0.367	0.354	0.194
50%	SNIP	<b>0.519</b>	0.413	<b>0.668</b>	0.276	<b>0.362</b>	<b>0.356</b>	0.207
50%	SNIP/ours(全距離)	0.516	<b>0.415</b>	0.698	<b>0.274</b>	0.384	0.375	<b>0.187</b>
50%	SNIP/ours(近距離)	0.513	0.410	0.703	0.275	0.395	0.390	<b>0.187</b>
70%	SNIP	0.491	0.386	0.720	0.284	0.493	0.448	0.198
70%	SNIP/ours(全距離)	<b>0.495</b>	<b>0.403</b>	<b>0.708</b>	<b>0.277</b>	<b>0.434</b>	<b>0.424</b>	<b>0.185</b>
70%	SNIP/ours(近距離)	0.487	0.381	0.736	<b>0.277</b>	0.450	0.470	0.198

表 2: 近距離における定量評価

枝刈り率	Methods	NDS	mAP	mATE	mASE	mAOE	mAVE	mAAE
0%	枝刈りなし	0.608	0.562	0.564	0.267	0.338	0.343	0.218
50%	SNIP	0.605	0.552	0.583	<b>0.265</b>	0.327	0.322	<b>0.212</b>
50%	SNIP/ours(全距離)	0.595	0.537	0.589	0.270	0.396	0.330	0.223
50%	SNIP/ours(近距離)	<b>0.623</b>	<b>0.579</b>	<b>0.565</b>	0.267	<b>0.288</b>	<b>0.275</b>	0.270
70%	SNIP	0.594	0.541	<b>0.564</b>	<b>0.267</b>	0.369	<b>0.356</b>	<b>0.201</b>
70%	SNIP/ours(全距離)	0.596	0.546	0.657	0.274	0.348	0.357	0.203
70%	SNIP/ours(近距離)	<b>0.598</b>	<b>0.548</b>	0.605	<b>0.267</b>	<b>0.347</b>	0.359	<b>0.201</b>

### 4.3. 定性評価

枝刈り率 50% と枝刈り率 70% の可視化画像結果を図 2 と図 3 に示す。図 2 と図 3 より、SNIP/ours(近距離) の赤枠で示した歩行者の検出結果より、精度低下が抑制されていることが確認できる。

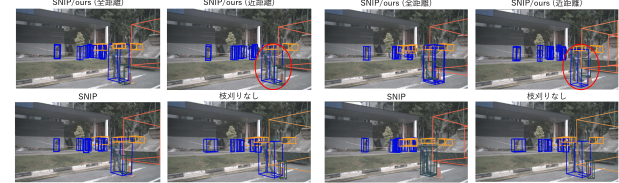


図 2: 50% 可視化画像

図 3: 70% 可視化画像

## 5. まとめ

本研究では、精度劣化を抑制した BEVFormer の重み付き非構造化枝刈り手法を提案した。評価実験より、全距離評価では SNIP/ours(全距離) が SNIP を上回り、近距離評価では SNIP/ours(近距離) が NDS/mAP を改善した。今後は自動運転における評価として提案手法が運転経路に与える影響を検証する。

## 参考文献

- [1] T. Ito *et al.*, “Weight Pruning to Mitigate Class-Specific Accuracy Degradation for LiDAR-Based 3D Object Detection,” in Proc. IEEE Intelligent Vehicles Symposium (IV), 2025.
- [2] Z. Li, *et al.*, “BEVFormer: Learning Bird’s-Eye-View Representation from Multi-Camera Images via Spatiotemporal Transformers,” In European Conference on Computer Vision (ECCV), 2022.