

## 1. はじめに

自動運転システムの実現には、点群データを用いて走行環境における物体（車両、歩行者、自転車等）を高速かつ正確に検出する 3D 物体検出タスクが重要である。高速性と高精度を同時に達成するには、性能を維持しつつモデルサイズを小さくする枝刈り等のモデル圧縮技術が重要となる。3D 物体検出タスクに対する枝刈りはこれまでに研究がされておらず、3D 物体検出モデルに既存の枝刈り手法を適用した際の傾向も不明である。

そこで本研究では、3D 物体検出モデルを既存手法 [1-4] で枝刈りした際の問題点を予備実験により明確にし、その問題点を解決する 3D 物体検出モデルに適した枝刈り手法を提案する。予備実験より、既存手法では特定クラスの検出能力が著しく低下することがわかった。そこで、特定クラスの知識を保持するように重みパラメータを評価する枝刈り手法を提案する。これにより、各クラスの精度低下の偏りを抑制できるようになる。

## 2. ニューラルネットワークの枝刈り

ニューラルネットワークの枝刈りは、冗長な重みパラメータを削除することでモデルを圧縮する技術である。2D 画像分類タスクを対象とした枝刈りの研究は盛んに行われており、多数の手法が提案されている。枝刈りのアプローチは、構造的枝刈りと非構造的枝刈りの 2 つに大別することができる。構造的枝刈りは層やフィルタ、重み行列の次元単位で重みを削除するため、推論の高速化が期待できる。非構造的枝刈りは特定の重み単位で重みを削除するため、構造的枝刈りと比べて性能を維持できる一方、GPU を用いた推論処理の高速化は困難である。しかしながら、非構造的枝刈りに対応した専用ハードウェアの研究も行われている。本研究では、高い性能を維持しつつ専用ハードウェアによる高速化が期待できる非構造的枝刈りを対象とする。

非構造的枝刈りの先行研究では、様々な観点で冗長な重みパラメータを特定する評価基準が提案されている [1-4]。Magnitude Magnitude [1] は、重みパラメータの絶対値が小さい順に重みを削除する手法である。Magnitude の重みパラメータ評価式  $S_{\text{Magnitude}}$  は、重みパラメータを  $w_i \in \mathbf{W}$  として式 (1) で導出する。

$$S_{\text{Magnitude}}(w_i) = |w_i| \quad (1)$$

Single-Shot Network Pruning Single-Shot Network Pruning (SNIP) [2] は、損失関数に影響を与える重みパラメータを削除する手法である。SNIP の重みパラメータ評価式  $S_{\text{SNIP}}$  は、損失関数を  $\mathcal{L}$ 、入力を  $\mathbf{x}$  として式 (2) で導出する。

$$S_{\text{SNIP}}(w_i) = \left| \frac{\partial \mathcal{L}(\mathbf{x})}{\partial w_i} w_i \right| \quad (2)$$

SNIP-Magnitude SNIP-Magnitude [3] は、値の大きい重みを保持しつつ、下流タスクに最適化される重みを残すように重みパラメータを削除する手法である。SNIP-Magnitude の重みパラメータ評価式  $S_{\text{SNIP-Mag}}$  は、ハイパーパラメータを  $\alpha$  として式 (3) で導出する。

$$S_{\text{SNIP-Mag}}(w_i) = \left| \frac{\partial \mathcal{L}(\mathbf{x})}{\partial w_i} w_i \right| + \alpha w_i^2 \quad (3)$$

ReFer Single-shot Foresight Pruning for Retaining Feature Representation (ReFer) [4] は、各層  $l$  の出力特徴量を維持するように重みパラメータを削除する手法である。ReFer の重みパラメータ評価式  $S_{\text{ReFer}}$  は、出力特徴量の L1 ノルムの総和を  $\|f^l\|_1$  として式 (4) で導出する。

$$S_{\text{ReFer}}(w_i) = \left| \frac{\partial \sum_{l=1}^L \|f^l\|_1}{\partial w_i} w_i \right| \quad (4)$$

## 3. 3D 物体検出タスクにおける枝刈りの調査

3D 物体検出タスクにおける枝刈りは、2D 画像分類タスクで提案された既存手法を単純に適用した際の傾向が不明であり、ほとんど研究されていない。そこで、予備実験として 3D 物体検出モデルを 2D 画像分類タスクで提案された既存手法で枝刈りする実験を行い問題点を明確にする。

予備実験 既存手法を用いて事前学習済みの PointPillars [5] を枝刈り率 90% で枝刈り後、事前学習時のデータセットでファインチューニングした際の各クラスに対する精度を図 1 に示す。図 1 より、全ての既存手法で Pedestrian（歩行者）の精度低下が大きく、既存手法による枝刈りでは特定のクラスの検出能力が著しく低下することがわかる。そのため、3D 物体検出モデルは各クラスの精度低下の偏りを抑制して枝刈りする必要があると考える。

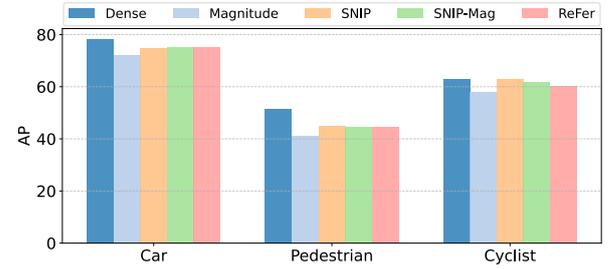


図 1: 既存手法で枝刈りした際の各クラスに対する精度

## 4. 提案手法

予備実験より、既存手法による枝刈りでは特定のクラスを検出する知識が失われることが判明した。そこで本研究では、特定の知識を保持する重みパラメータ評価基準を提案する。具体的には、知識を保持したいクラスに対して有効な重みパラメータを SNIP を用いて特定する。これにより、特定のクラスの検出に有効な重みパラメータを残すように枝刈りすることで、そのクラスの精度低下の抑制を期待する。

提案手法による重みパラメータの評価式  $S_{\text{ours}}$  は、点群データ  $\mathbf{P}$  に含まれる知識を保持したいクラスの  $n$  個のサンプル  $\{\mathbf{x}_1^{\text{Specific}}, \dots, \mathbf{x}_n^{\text{Specific}}\} \subset \mathbf{P}$  をモデルに入力した際の損失関数を  $\mathcal{L}(\mathbf{x}_{1:n}^{\text{Specific}})$  をとして、式 (5) で表現できる。

$$S_{\text{ours}}(w_i) = \left| \frac{\partial \mathcal{L}(\mathbf{x}_{1:n}^{\text{Specific}})}{\partial w_i} w_i \right| \quad (5)$$

SNIP は入力に対する損失を維持することでその知識を保持できるため、入力を知識を保持したいクラスのサンプルとすることで特定の知識を保持できることを期待する。知識を保持したいクラスのサンプルは事前学習済みモデルにより検出が成功しているサンプルである必要がある。また、提案手法は全てのクラスを検出するための知識の保持は考慮されていないため、全体的な精度を維持するように枝刈りする既存手法による評価も同時に考慮する必要がある。提案手法による最終的な重みパラメータ評価式は、任意の既存手法を  $S_{\text{existing}}(w_i)$  として式 (6) で表現できる。

$$S(w_i) = S_{\text{ours}}(w_i) + S_{\text{existing}}(w_i) \quad (6)$$

## 5. 評価実験

本実験では、既存手法では大幅に精度が低下する特定クラスの検出に有効な重みを残すように枝刈りし、提案手法の有効性を検証する。

### 5.1 KITTI dataset

KITTI dataset は、LiDAR センサ等を自動車に取り付け、走行環境を点群データとして収集したデータセットである。Car, Pedestrian, Cyclist の 3 クラスで構成されて

おり, Easy, Moderate, Hard の 3 段階の難易度がある. 難易度が上がるほどサイズの小さい物体, 遮蔽物により視認性が低い物体が含まれる. 学習データは 3,712 枚, 評価データは 3,769 枚である.

## 5.2 実験概要

本実験では, KITTI dataset で事前学習された PointPillars を用いて枝刈りした後, KITTI dataset でファインチューニングをし評価を行う. PointPillars における知識を保持する特定クラスは, Pedestrian とする. 事前学習済み PointPillars は, 3D 物体検出タスクのためのツールである OpenPCDet で公開されている重みパラメータを用いる. 枝刈り率は 90% とし, 提案手法と併用する既存手法  $S_{existing}(w_i)$  には Magnitude, SNIP, SNIP-Magnitude, ReFer を用いる. ファインチューニング時のエポック数を 30, バッチサイズを 64, 最適化手法に Adam を用い, 学習率は  $3e-4$ , Weight decay は 0.01 とする.

評価指標には, Intersection over Union (IoU), Average Precision (AP), mean Average Precision (mAP) を用いる. IoU は, 予測バウンディングボックスと正解バウンディングボックスの重なり具合を表す. 本実験で使用する各クラスの IoU の閾値は Car が 70%, Pedestrian と Cyclist が 50% であり, 難易度は Moderate とする. AP は, 縦軸を再現率, 横軸を適合率とした際の曲線の面積を表す. mAP は, 各クラスの AP を平均した値である.

## 5.3 実験結果

枝刈り率 90% 時の各枝刈り手法における mAP, 各クラスの AP を表 1 に示す. Dense は枝刈り前の精度である. 表 1 より, 提案手法と併用することにより各既存手法の mAP が向上することが確認できる. また, ReFer を除いて Pedestrian の精度低下を抑制できることを確認した.

次に, 枝刈りされる層の傾向を確認する. 各枝刈り手法における層ごとの Sparsity を図 2 に示す. Sparsity の値が高いほど枝刈りされていることを示し, グラフの背景色は図 3 に示す PointPillars のモデル構造における各モジュールの色に対応する. 図 2 より, 紫色で示す提案手法は畳み込み層の浅い層 (conv1) で既存手法よりも低い Sparsity であることが確認できる. これは, 浅い層は解像度が高い特徴マップを扱うことから, 提案手法が Pedestrian のような小さな物体を検出する知識の保持に有効であると評価したためであると考えられる. 以上より, 提案手法により既存手法では削除されてしまう特定の知識を保持できることがわかる.

表 1: 各既存手法及び提案手法との精度比較

	mAP	Car	Pedestrian	Cyclist
Dense	64.25	78.38	51.45	62.93
Magnitude	57.07	72.08	41.26	57.86
Ours w/ Magnitude	<b>60.08</b>	<b>74.10</b>	<b>45.11</b>	<b>61.04</b>
SNIP	60.84	<b>74.91</b>	44.81	62.81
Ours w/ SNIP	<b>62.21</b>	74.55	<b>48.46</b>	<b>63.62</b>
SNIP-Mag	60.48	<b>75.15</b>	44.43	61.86
Ours w/ SNIP-Mag.	<b>61.14</b>	75.02	<b>46.21</b>	<b>62.19</b>
ReFer	60.02	<b>75.19</b>	<b>44.43</b>	60.44
Ours w/ ReFer	<b>60.20</b>	74.33	43.82	<b>62.45</b>

## 5.4 提案手法の汎用性検証

本実験では, 他のモデルである SECOND 及び PV-RCNN を用いて提案手法の汎用性を検証する. 実験条件は, 枝刈り率を 95%, PV-RCNN の学習率を  $1e-2$ , PV-RCNN のバッチサイズを 16 とし, その他の条件は 5.2 節に準拠する. SECOND 及び PV-RCNN における知識を保持する特定クラスは, 既存手法では失われてしまう Car とする. 提案手法と併用する既存手法は, 既存手法の中で最も性能低下を抑制していた手法とする. SECOND 及び PV-RCNN における mAP, 各クラスの AP を表 2 に示す. 表 2 より,

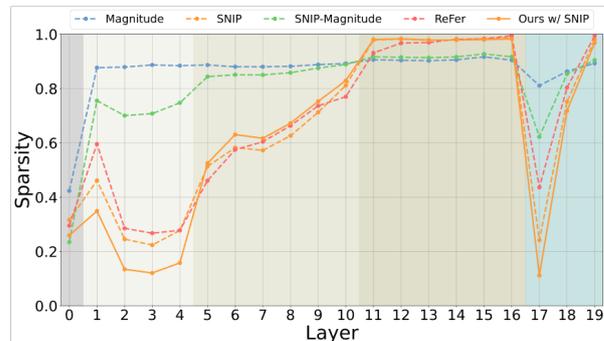


図 2: 各枝刈り手法における層ごとの Sparsity



図 3: PointPillars のモデル構造

提案手法と併用することにより Car の精度低下を抑制できることを確認した.

表 2: SECOND 及び PV-RCNN における精度比較

Method	SECOND			
	mAP	Car	Pedestrian	Cyclist
Dense	66.49	81.60	51.15	66.74
SNIP	<b>62.61</b>	75.27	<b>48.09</b>	<b>64.46</b>
Ours w/ SNIP	61.44	<b>77.68</b>	43.06	63.58
Method	PV-RCNN			
	mAP	Car	Pedestrian	Cyclist
Dense	69.73	84.36	54.46	70.38
SNIP-Magnitude	67.47	81.48	<b>53.06</b>	67.88
Ours w/ SNIP-Mag.	<b>69.13</b>	<b>82.50</b>	51.38	<b>73.51</b>

## 6. おわりに

本研究では, 3D 物体検出モデルにおける枝刈りの問題に対処する知識を保持する枝刈り手法を提案した. 評価実験より, 提案手法は特定クラスの精度低下を抑制できることを確認した. 今後は, LiDAR センサから遠距離に位置する物体など, 他の知識を保持する実験を行い, 提案手法の有効性を検証する予定である.

## 参考文献

- [1] S. Han, *et al.*, “Learning both Weights and Connections for Efficient Neural Networks”, In ACM, 2015.
- [2] N. Lee, *et al.*, “SNIP: SINGLE-SHOT NETWORK PRUNING BASED ON CONNECTION SENSITIVITY”, In ICLR, 2019.
- [3] H. Kohama, *et al.*, “Single-Shot Pruning for Pre-trained Models: Rethinking the Importance of Magnitude Pruning”, In ICCVW, 2023.
- [4] 新田常顧 等, “事前学習モデルの特徴表現を維持した Single-shot Foresight Pruning”, 画像の認識・理解シンポジウム, 2024.
- [5] A. Lang, *et al.*, “PointPillars: Fast Encoders for Object Detection from Point Clouds”, In CVPR, 2019.

## 研究業績

- [1] T. Ito, *et al.*, “Analyzing the Accuracy, Representations, and Explainability of Various Loss Functions for Deep Learning”, International Joint Conference on Neural Networks (IJCNN), 2023.
  - [2] 伊藤天詞 等, “画像分類タスクにおける損失関数と注視領域の傾向調査”, 画像の認識・理解シンポジウム, 2023.
- (他 4 件)