



## 最新論文調査：3D Object Detection編

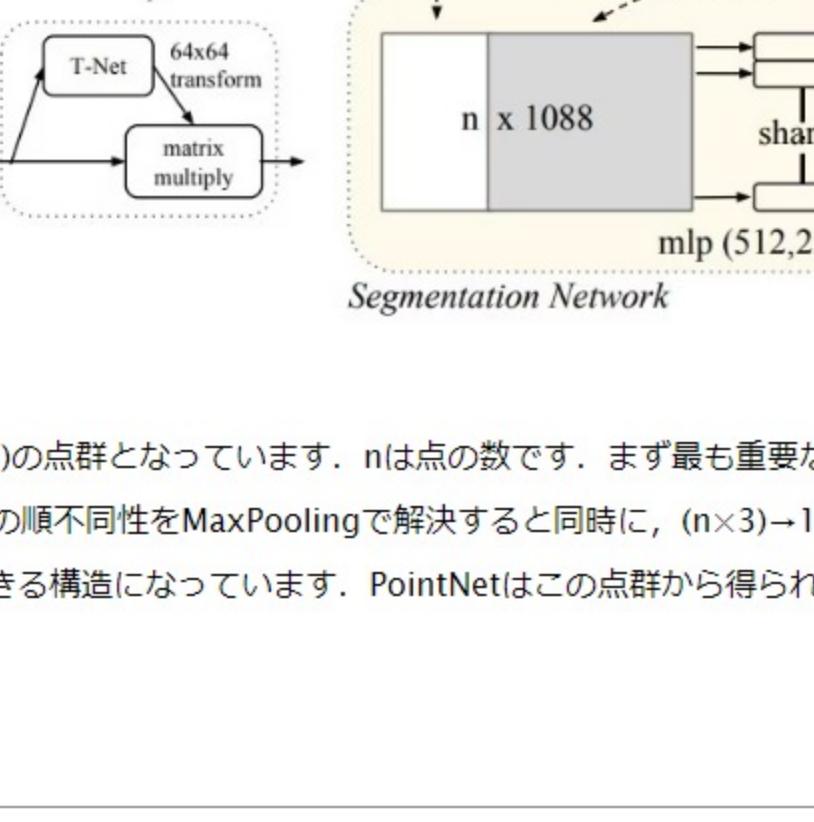
2023-04-07 17:00 ▲ AI技術開発部

AI Tech



## 3D Object Detectionとは

カメラやLiDARというセンサーの情報を入力として3次元座標におけるbounding box(bbox)を予測するタスクです。下図では、LiDARが反射した位置に点が打たれています。実際に予測するのは、基本は7次元の(x, y, z, l, w, h, yaw)です。(x, y, z)はbbox中心の座標です。(l, w, h)は各xyzの長さです、yawはヨー角、z軸回転のこと、[-π,+π]で対象の物体の角度を表します。一般的車はウイリーのような縦横に回転することはないため、z軸における角度のみを回転することで、車の向きを特定します。



[1]より引用

3D Object Detectionのタスクでは、入力の多くはLiDARというセンサーから取られた点群です。点群は画像と異なり、距離の情報が取得できます。そのため、距離の予測が画像と比べて正確にできます。また、画像とLiDARを結合させて予測するモデルもありますが、多くのセンサフュージョンアプローチでは推論速度が大幅に遅いといった問題があり、LiDAR単体によるアプローチが主流となっています。

## 論文紹介

## PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation (2017) [2]

PointNetは点群固有の入力の問題を解決した論文です。PointNet自体は物体検出ではなく、点群全体のクラス分類やセグメンテーション手法として提案されていますが、その特徴抽出の部分は今後の点群を入力とした研究において広く用いられています。

点群を入力とするネットワークの場合、画像と異なり以下の特徴的な問題があります。

「順不同性」—入力の点の順番を入れ替えると同じ出力にならなければなりません」という問題です。

PointNetでは順不同性を解決する手法が「MaxPooling」であるとし、その有効性を示しました。



図はPointNetの構造です。入力は(n×3)の点群となっています。nは点の数です。まず最も重要なのは、global featureです。1024は任意の特徴量次元ですが、PointNetは点群固有の順不同性をMaxPoolingで解決すると同時に、(n×3)→1024となり、データによって異なる点群数nにようらず一定のサイズの特徴量を出力できる構造になっています。PointNetはこの点群から得られたglobal featureをもとに点群のクラス分類やセグメンテーションを行っています。

## PointPillars: Fast Encoders for Object Detection from Point Clouds (2019) [3]

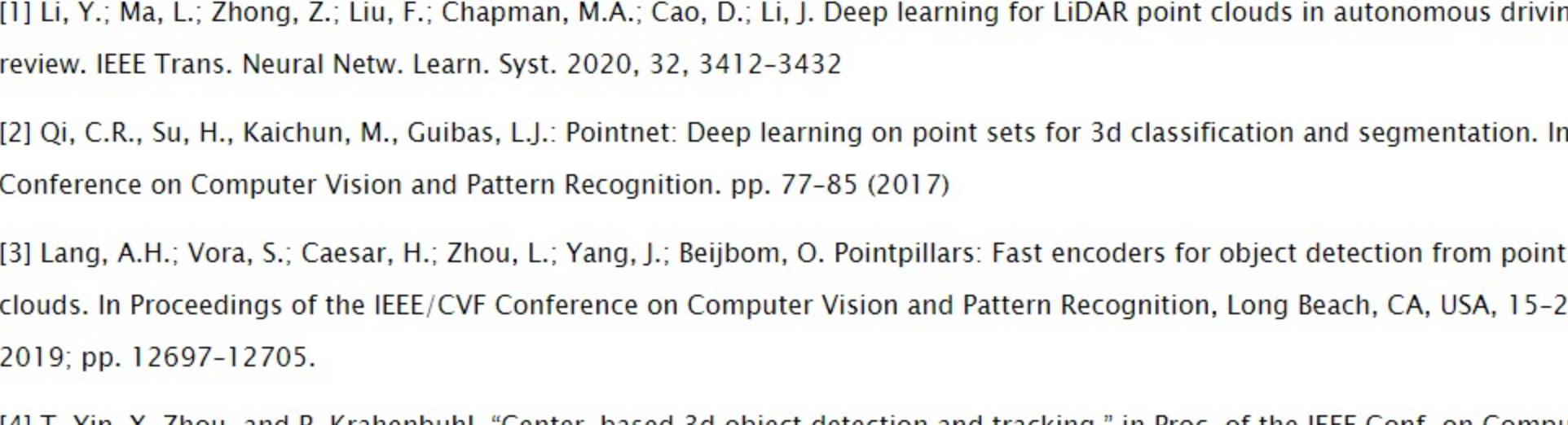
PointNetを用いた自動運転3次元物体検出のネットワークであり、超高速かつ高い精度であることから、今後さまざまなモデルのベースラインとして用いられています。



PointPillarsでは大きく分けると3つのページに分けられます。Pillar Feature Netでは、領域をPillar(柱)に分割し、点群をいずれかのPillarに所属させます。この各Pillarの中に入れた点をPointNetに入力し、それぞれのPillarから特徴量を抽出します。総のPillar数H×横のPillar数W×特徴量Cとなるため、データの形状は(H, W, C)となります。これは画像と同じ形のデータになります(Pseudo image)。これを画像とみなすことで、研究が盛んに行われている既存の画像用物体検出器(例Single Shot Detector: SSD)を使えるようにしました。Backboneでは、画像と同様のCNNを用いた畳み込みで特徴量を抽出しています。Detection Headでは、主に(dx, dy, z, l, W, H, yaw)の7次元のラメータの回帰とクラス分類を行っています。SSDはアンカーベース手法であるため、各アンカー中心との差(dx, dy), 物体の高さz, bboxのサイズLWH, bboxの角度yawを回帰します。

## Center-based 3D Object Detection and Tracking (2021) [4]

CenterPointはPointPillarsのheadを改良したモデルで、2Dの物体検出器のCenterNetから着想を得ています。こういった3次元物体検出におけるデータの特徴として、道路の質感上、周辺の車は自分と同じ向きが真逆(0度or180度回転)であることが多いです。自分の位置から生成されるアンカーベースの手法では、自分がカーブの轍中であるなど、自分の車周りの車の角度が並行ではないときに、角度の予測を大きく誤ることに注目しました。



PointPillarsではこういった内容を解決するために、アンカーベースな物体検出を行います。下図はCenterPointのネットワークの概要です。3D Backboneにあたる特徴抽出器はPointPillarsと同様の構造で画像用の特徴量を抽出します。

Headにあるのは、SSDに代わりCenterNetのような構造をしています。まず、Birds Eye View(BEV)での、物体の中心座標のみをヒートマップで予測します。このヒートマップを用いた物体検出器はCenterNetのアイディアです。各物体の中心座標だけをアンカーを用いてヒートマップで予測した後、ヒートマップの特徴量から各パラメータ(xy偏差、サイズや角度)をさらに回帰します。この2stage構造することで、角度の推論を向上させ、NuScenesデータセットにおける物体検出とトラッキングの両方でSoTAを達成したことが報告されています。



## 再現実装

3D Object Detectionのデータセットでは、KITTIやWaymoが有名ですが、今回はPandaSet(CC BY 4.0)というデータセットを用います。PandaSetは米ScaleAI社が提供するサンフランシスコでの自動運転データセットです。今回はCenterPointをPandaSetに対応させ、学習、デモを行いました。実際のテストデータにおける出力が以下の図です。



## おわりに

3D Object Detectionについての論文をいくつか紹介させていただきました。3D Object Detectionは自動運転のビジョンとなる研究です。この研究が発展することで完全自動運転が実現する日が早くなるでしょう。今後の発展、実用化が大いに期待されます。本記事が皆様にとって有益な情報であれば、幸いです。

## 引用

[1] Li, Y.; Ma, L.; Zhong, Z.; Liu, F.; Chapman, M.A.; Cao, D.; Li, J. Deep learning for LiDAR point clouds in autonomous driving: A review. *IEEE Trans. Neural Netw. Learn. Syst.* 2020, 32, 3412–3432.

[2] Qi, C.R.; Su, H.; Kaichun, M.; Guibas, L.J. PointNet: Deep learning on point sets for 3d classification and segmentation. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 77–85 (2017)

[3] Lang, A.H.; Vora, S.; Caesar, H.; Zhou, L.; Yang, J.; Beijbom, O. Pointpillars: Fast encoders for object detection from point clouds. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 15–20 June 2019; pp. 12697–12705.

[4] T. Yin, X. Zhou, and P. Krahenbuhl, "Center-based 3d object detection and tracking," in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 2021, pp. 11 784–11 793.

## AI技術開発部

人行動、医療画像、外観検査など多用な機械学習モデルの開発に取り組んでいます。

## 前の記事



現地施工員がFORXAIで温湿度・CO2を見える化してみた

## 次の記事



データサイエンス導入での産官学連携の取り組み～学術の知見をビジネスに社会実装する！～

シェアする

ポスト

ブックマーク

Pocket

LINEで送る

Emailで送る

2020-2023 Konica Minolta, Inc.