

進捗報告

1 今週やったこと

- M2Det の調査
- PointPillars の調査
- SSD の実装

2 今週の収穫

2.1 M2Det の調査

前回のゼミで SSD の改良版として RefineDet を紹介したが, その更なる改良版として M2Det [1] というものがあつたのでそれについて調べてみた.

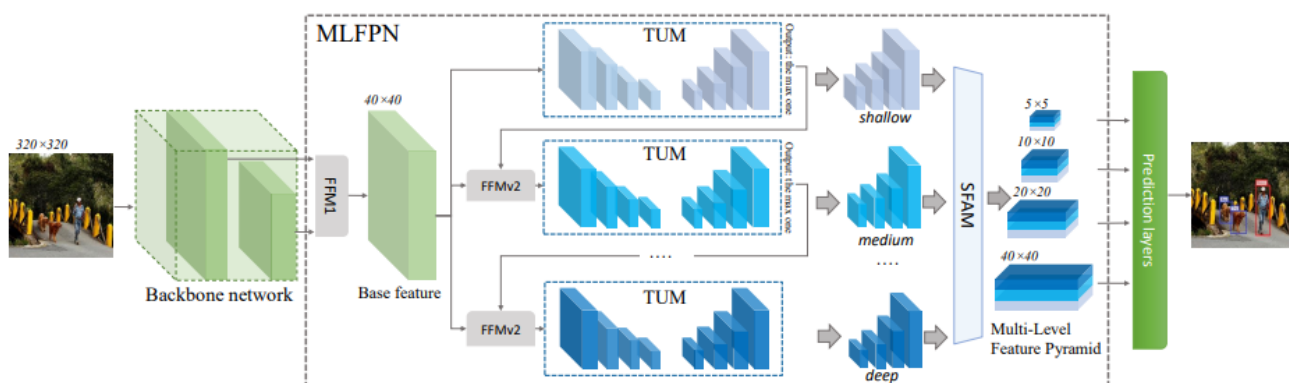


図 1: M2Det のアーキテクチャ概要図 (論文 [1] 図 2 引用)

図 3 に M2Det のアーキテクチャ概要図を示す. M2Det は大きく分けて Backbone network, MLFPN (Multi-Level Feature Pyramid Network), Prediction layers の 3 つで構成されている.

Backbone network は単なる特徴抽出器であり, VGGNet や ResNet などの任意のアーキテクチャが利用できる. もしかするとここに PointNet を使うこともできるかもしれない.

MLFPN は FPN というネットワークの構造を受け継いだ上で FPN の構造を Multi-Level にしたものである. 図 2 に FPN のアーキテクチャ概要図を示す. この構造により低い層の特徴も利用することができ, 解像度が高くなる.

また MLFPN では次の 3 つのモジュールで構成されている.

- FFM: 異なる解像度の特徴マップをフュージョンすることで Base feature と呼ぶ特徴マップを生成する.
- TUM: エンコーダ-デコーダ構成となっており, FFM からの出力を受け取って再度マルチスケールな特徴ピラミッドを生成する.
- SFAM: 多段に接続した各 TUM から得られる特徴ピラミッドを最後に統合する.

この構造により, 特に小さな物体の検出に関して M2Det は SSD や YOLO を凌駕する結果を残している. なお小さな物体の検出が果たして星座の検出に有利に働くかは不明である.

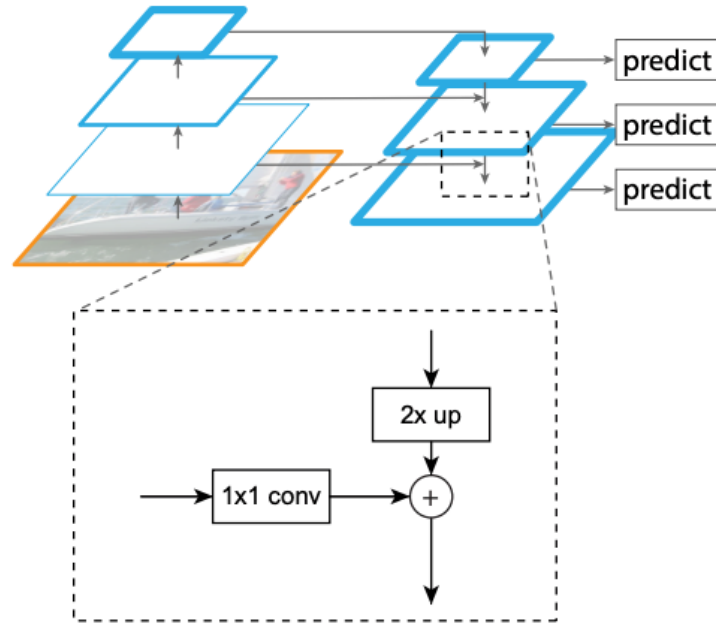


図 2: FPN のアーキテクチャ概要図 (論文 [2] 図 3 引用)

2.2 PointPillars の調査

PointPillars [3] は PointNet の構造を基にして点群から物体検出をするネットワークである. 図 ?? に PointPillars のアーキテクチャ概要図を示す.

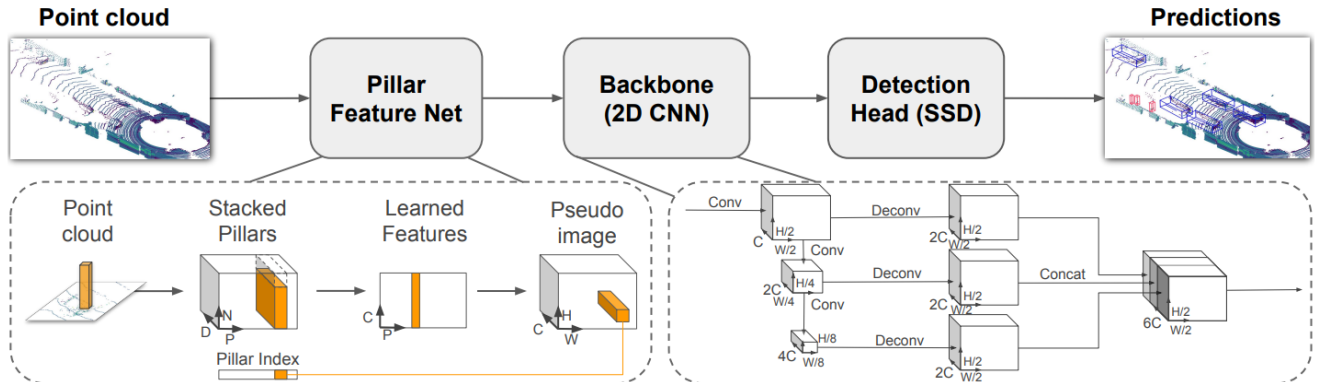


図 3: PointPillars のアーキテクチャ概要図 (論文 [3] 図 2 引用)

PointPillars は次の 3 つのネットワークで構成されている.

- Pillar Feature Net: 点群をスパースな pseudo image に変換する.
- Backbone (2D CNN): pseudo image をハイレベル特徴量へ変換する.
- Detection Head(SSD): 検出と 3D ボックスへの回帰を行う.

PointPillars はこのように 3D 点群を柱に分割し, 2D のフレームワークに畳み込むことで高速化を実現している. また性能としても比較的高い精度を出している.

2.3 SSD の実装

M2Det にせよ PointPillars にせよ比較対象が必要であるため、まず基礎実験として単純な SSD を実装して結果を確認してみようと考えている。実装は現在している最中である。なおデータセットに関しては各星座の上位 8 個を抽出した画像に関してアノテートしたものを作成した。図 4 にアノテートの一例を示す。訓練用として、1200 枚の星座画像に対してアノテート作業をした。



図 4: アノテートの一例

3 今後の方針

SSD を動かしてみて、まずは結果を確認したいと考えている。

参考文献

- [1] Zhao, Q., Sheng, T., Wang, Y., Tang, Z., Chen, Y., Cai, L., & Ling, H. (2019). M2Det: A Single-Shot Object Detector Based on Multi-Level Feature Pyramid Network. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 9259-9266.
- [2] T. Lin, et al., "Feature Pyramid Networks for Object Detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017 pp. 936-944.
- [3] Lang, A.H., Vora, S., Caesar, H., Zhou, L., Yang, J., & Beijbom, O. (2019). PointPillars: Fast Encoders for Object Detection From Point Clouds. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 12689-12697.