

進捗報告

1 今週やったこと

- SSD の改変方法の検討

2 今週の収穫

2.1 SSD の改変方法の検討 その 1

SSD の検出結果から, 星同士の位置関係が不変である良さを活かせていないことがわかったため, 損失関数にその要素も組み込むことを検討している. 以下に SSD の損失関数を示す.

$$\begin{aligned}
 L(x, c, l, g) &= \frac{1}{N} (L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \\
 L_{conf}(x, c) &= - \sum_{i \in Pos} x_{ij}^p \log \hat{c}_i^p - \sum_{i \in Neg} x_{ij}^p \log \hat{c}_i^0 \quad \text{where} \quad \hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \\
 L_{loc}(x, l, g) &= \sum_{i \in Pos} \sum_{m \in cx, cy, w, h} x_{ij}^k \text{smooth}_{L1}(l_i^m - \hat{g}_j^m) \quad \text{where} \quad \text{smooth}_{L1} = \begin{cases} 0.5x^2 & (|x| < 1) \\ |x| - 0.5 & (\text{otherwise}) \end{cases}
 \end{aligned}$$

SSD の損失関数は確信度誤差である L_{conf} と位置特定誤差である L_{loc} の和で, α はハイパーパラメータである. ここで, x はカテゴリ p において, i 番目のデフォルトボックスと j 番目の正解ボックスの一致度を示すものであり, 一致すれば 1, 一致しなければ 0 となる. また N はマッチしたデフォルトボックスの数, c はデフォルトボックスの座標のオフセット, l は予測されたボックス, g は正解ボックスである.

この損失関数に, 星座の位置関係に関する項を追加したい.

現状, SSD を改変することで 8732 個のデフォルトボックスの中で, 物体を発見したデフォルトボックスの位置情報とそのラベルを獲得することができている. 図 1 に獲得した情報の具体例を示す.

なおこの情報は 1 バッチ (今回はバッチサイズが 6) まとめての情報である. この情報から各クラスのデフォルトボックスの中心間距離を算出することができるため, この情報をうまく扱うことで各クラス間の距離を考慮した学習ができるようになると考えている. ここで考えるべきなのが, すべての組み合わせにおいてクラス間距離を求めるのがいいのか, クラス集合の中心間距離を求めるのがいいのか, ということである. この算出距離と正しい距離との差分をとった値 (個人的には位置特定誤差でも使用されている Smooth L1 Loss を使用したいと考えている) を損失関数に追加したいと考えている.

2.2 SSD の改変方法の検討 その 2

もう一つのクラス間距離を考慮する方法として, 相対的な位置関係を考慮した一般物体検出手法を提案している論文 [1] が参考になると考えている. これはファッションの研究において, トップスやボトムス, シューズなどを検出する際にトップスやシューズは検出されているのにボトムスは検出されない, というような問題を回避するためにトップス, ボトムス, シューズの 3 つをセットで検出させようとする手法である. 図 2 に論文内に記載されているネットワーク構造を示す.

ただ論文内の情報だけではネットワークの理解をしきれずなかなか実装が難しそうではある.

▼デフォルトボックスの位置情報 (cx, cy, w, h)

```
tensor([
  [0.5733, 0.7600, 0.1414, 0.1414], [0.6000, 0.7600, 0.1414, 0.1414], [0.6267, 0.7600, 0.1414, 0.1414], [0.5733, 0.7867, 0.1414, 0.1414],
  [0.6000, 0.7867, 0.1414, 0.1414], [0.6267, 0.7867, 0.1414, 0.1414], [0.6133, 0.7200, 0.2000, 0.2000], [0.5600, 0.7733, 0.2000, 0.2000],
  [0.5600, 0.7733, 0.2828, 0.1414], [0.6133, 0.7733, 0.2000, 0.2000], [0.6133, 0.7733, 0.2828, 0.1414], [0.6667, 0.7733, 0.2828, 0.1414],
  [0.3200, 0.3200, 0.6192, 0.6192], [0.3200, 0.5333, 0.6192, 0.6192], [0.3200, 0.7467, 0.6192, 0.6192], [0.1667, 0.5000, 0.7100, 0.7100],
  [0.1667, 0.5000, 0.7904, 0.7904], [0.1667, 0.5000, 0.5020, 1.0000], [0.5000, 0.5000, 0.8800, 0.8800], [0.5000, 0.5000, 0.9612, 0.9612],
  [0.5000, 0.5000, 0.6223, 1.0000], [0.5333, 0.3200, 0.7637, 0.3818], [0.5333, 0.3200, 0.9353, 0.3118], [0.5000, 0.5000, 1.0000, 0.5020],
  [0.8267, 0.5600, 0.2720, 0.2720], [0.8800, 0.5600, 0.2720, 0.2720], [0.8267, 0.6133, 0.2720, 0.2720], [0.8800, 0.6133, 0.2720, 0.2720],
  [0.8267, 0.6667, 0.2720, 0.2720], [0.8800, 0.6667, 0.2720, 0.2720], [0.8267, 0.7200, 0.2720, 0.2720], [0.8800, 0.7200, 0.2720, 0.2720],
  [0.8000, 0.5867, 0.3700, 0.3700], [0.8000, 0.5867, 0.4470, 0.4470], [0.8000, 0.5867, 0.2616, 0.5233], [0.8000, 0.5867, 0.2136, 0.6409],
  [0.9067, 0.5867, 0.3700, 0.3700], [0.9067, 0.5867, 0.4470, 0.4470], [0.9067, 0.5867, 0.2616, 0.5233], [0.8000, 0.6933, 0.3700, 0.3700],
  [0.8000, 0.6933, 0.4470, 0.4470], [0.8000, 0.6933, 0.2616, 0.5233], [0.8000, 0.6933, 0.2136, 0.6409], [0.9067, 0.6933, 0.3700, 0.3700],
  [0.9067, 0.6933, 0.4470, 0.4470], [0.9067, 0.6933, 0.2616, 0.5233], [0.2667, 0.5867, 0.4470, 0.4470], [0.3733, 0.5867, 0.4470, 0.4470],
  [0.3200, 0.5333, 0.5400, 0.5400], [0.3200, 0.5333, 0.6192, 0.6192], [0.3200, 0.5333, 0.7637, 0.3818], [0.3200, 0.7467, 0.6192, 0.6192],
  [0.1667, 0.5000, 0.7100, 0.7100], [0.5067, 0.5067, 0.2720, 0.2720], [0.5067, 0.5067, 0.2828, 0.1414], [0.5600, 0.5067, 0.2720, 0.2720],
  [0.5600, 0.5067, 0.2828, 0.1414], [0.4533, 0.5600, 0.2000, 0.2000], [0.4533, 0.5600, 0.2720, 0.2720], [0.5067, 0.5600, 0.2000, 0.2000],
  [0.5067, 0.5600, 0.2720, 0.2720], [0.5067, 0.5600, 0.2828, 0.1414], [0.5600, 0.5600, 0.2000, 0.2000], [0.5600, 0.5600, 0.2720, 0.2720],
  [0.5600, 0.5600, 0.2828, 0.1414]], device='cuda:0')
```

▼デフォルトボックスのラベル情報

```
tensor([3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3,
3, 3, 3, 3, 3, 3, 3], device='cuda:0')
```

図 1: 獲得した情報の具体例

3 今後の方針

SSD を改変し, 実験を回す

参考文献

- [1] *佐々木 雄一, 相対的な位置関係を考慮した一般物体検出の改良, 人工知能学会全国大会論文集, 2019, JSAI2019 巻, 第 33 回 (2019), セッション ID 2M5-J-10-01, p. 2M5J1001, 公開日 2019/06/01, <https://doi.org/10.11517/pjsai.JSAI2019.0.2M5J1001>, https://www.jstage.jst.go.jp/article/pjsai/JSAI2019/0/JSAI2019.2M5J1001/_article/-char/ja

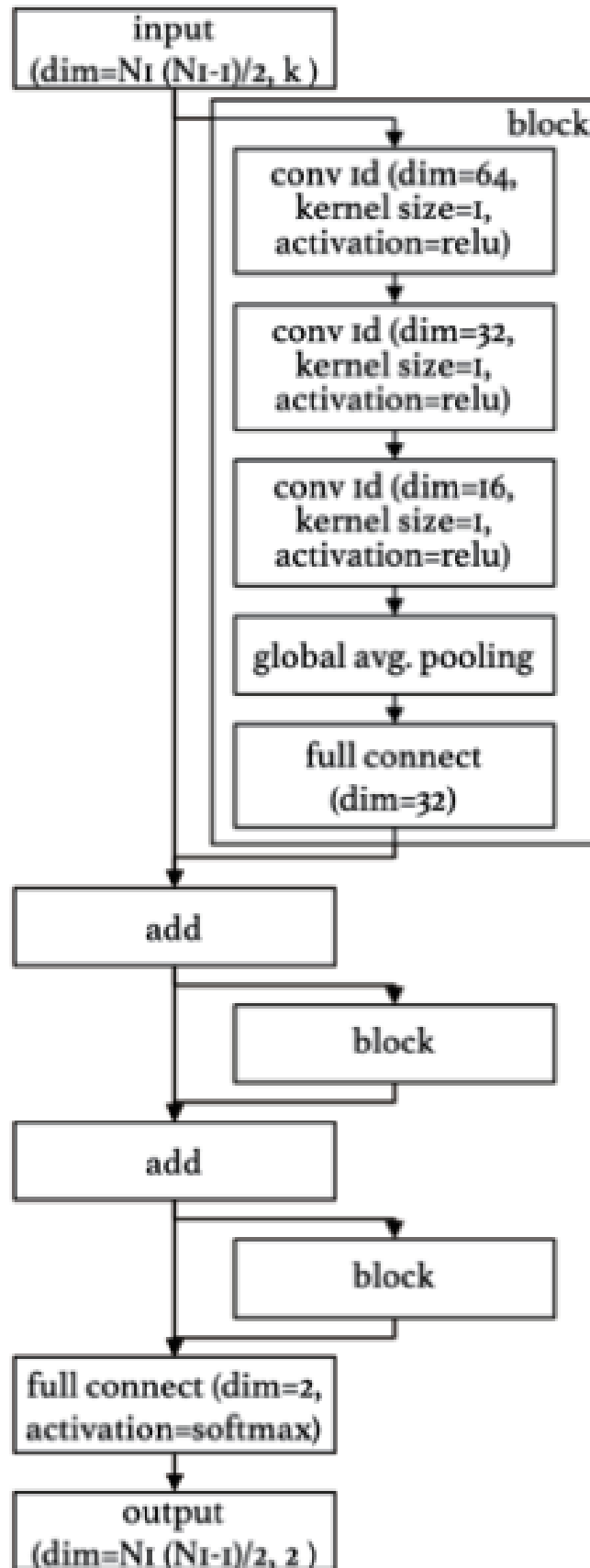


図 2: ネットワーク構造 (論文 [1] 図 2 より引用)