

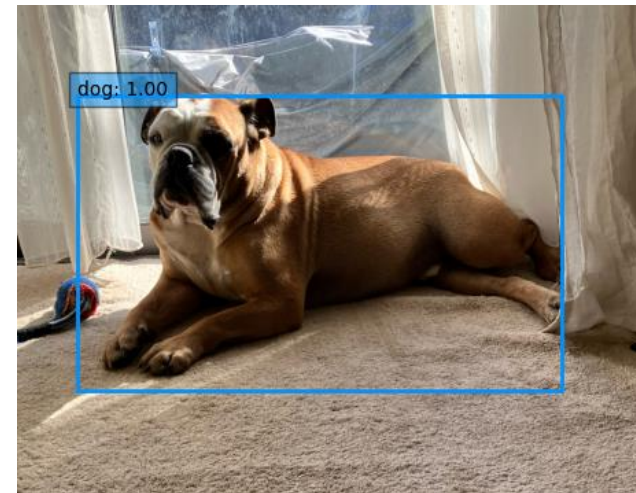
活動報告

(SSDモデルによる物体検出)

物体検出とは

以下のように、ある物体がどのクラスに属するかだけでなく、その物体の場所を特定する。

物体検出は一つの画像から複数の物体を特定することができる。
画像中の枠をBounding Box(Bbox)と呼ぶ。



有名なモデル

- SSD(Single Shot MultiBox Detector)

バランスが良い??

SSDの他にも、、、、、、

- YOLO(You Only Look Once)

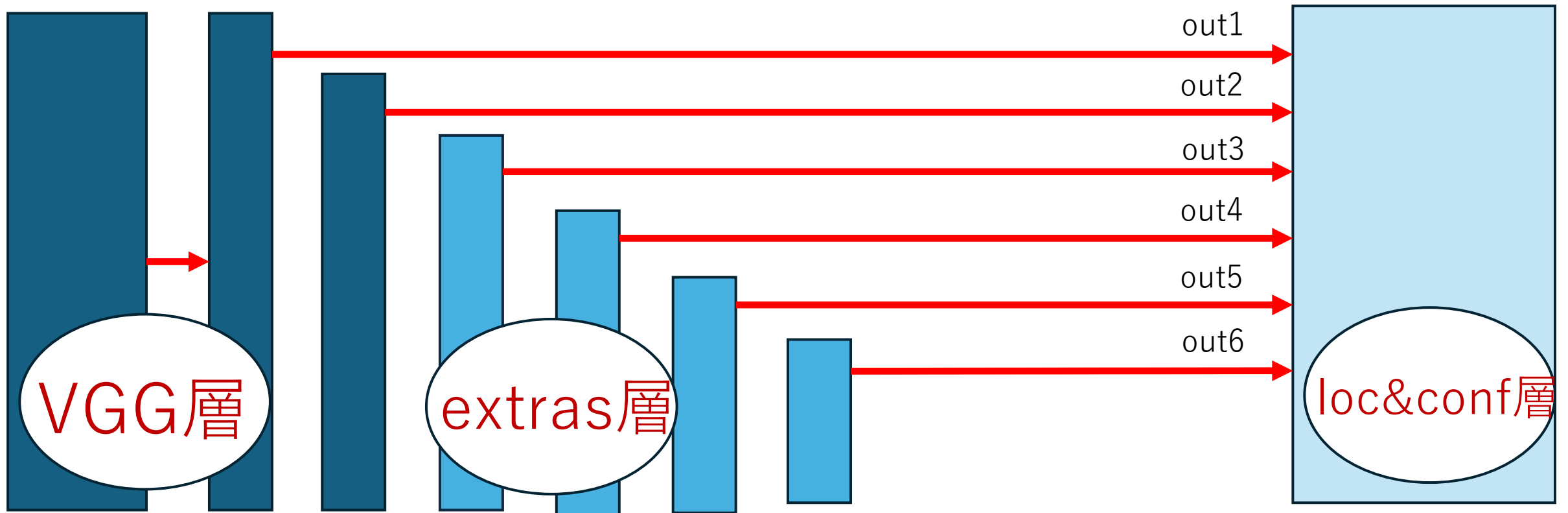
SSDよりも高速

- Faster R-CNN(R-CNN→Fast R-CNN →Faster R-CNN)

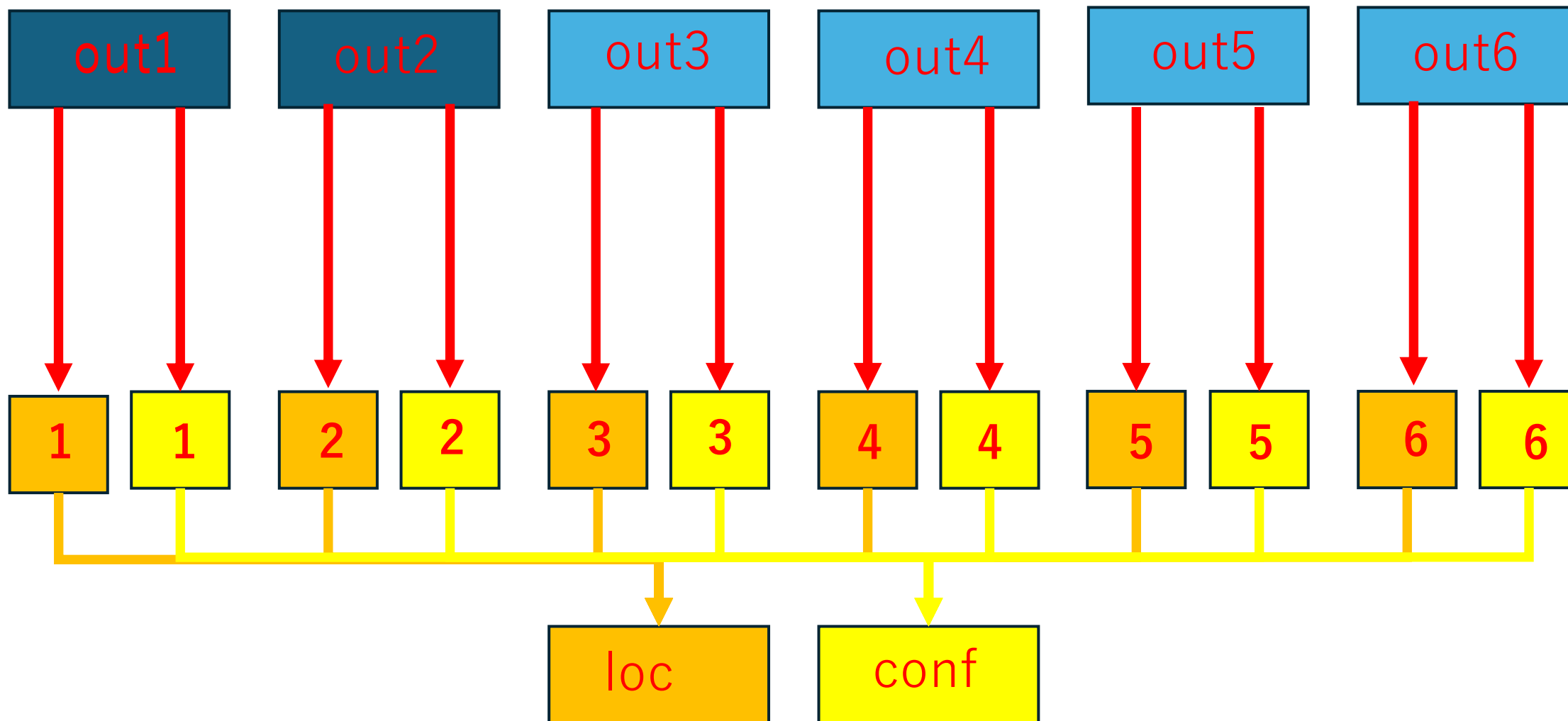
SSDよりも高精度

SSDの構造

- 以下のような構造にすることで、小さい物体と大きい物体を検出できるようにしている。



SSDの構造 (locとconf層の中身)



学習と推論の流れ

- ・ 学習

- ①データの前処理をする。
- ②各特徴マップの各セルごとにいくつかのデフォルトボックス (Dbox)を作る。
- ③前処理したデータをネットワークに入れる。Dboxごとにloc (オフセット値) とconfを求める。
- ④教師データとネットワークの出力値を比較して損失値を出す。
- ⑤損失値が減るようにパラメータを更新する。
- ⑥以上をエポック数分繰り返す。

Dboxとは

Bboxの元となるbox。右図のように一つのセルにつき、いくつかの形のDboxがあり、特徴マップを埋め尽くすように配置されている。

すべてのDboxについてlocとconfを求めて、locを使ってBboxにして、それぞれの物体につきconfをもとに一つのBboxを決定する。



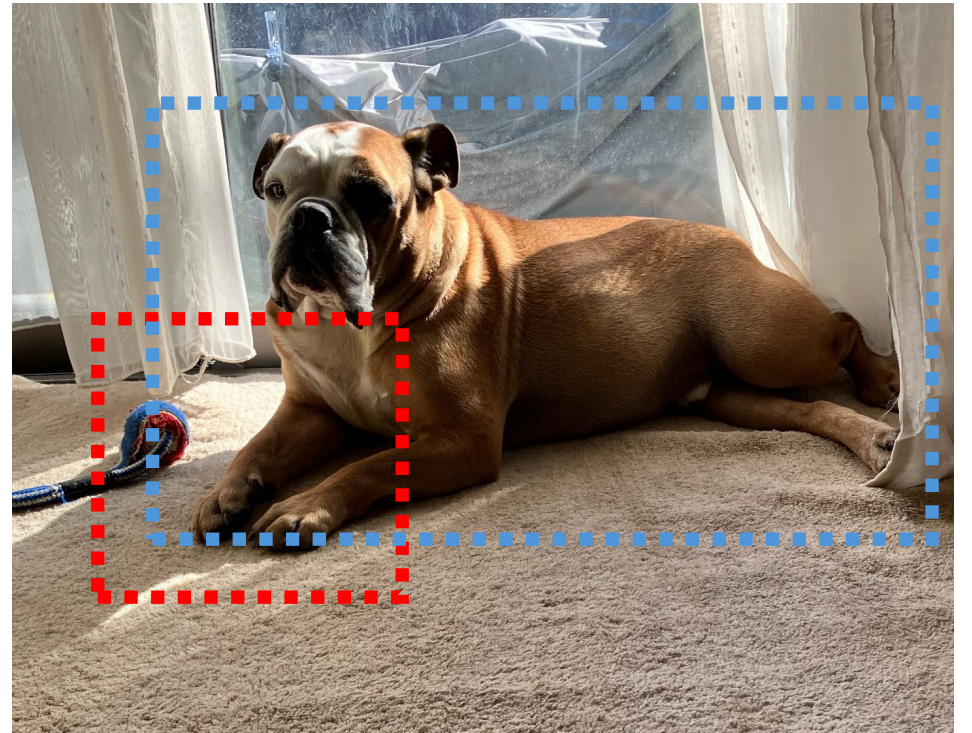
教師データの作成

各画像のBboxの座標をそのまま教師データには使用しない。それぞれのDboxがIoU値をもとに最もマッチするBboxをそれぞれの画像で求める。そして、そのBboxのクラスとオフセット値（DboxとBboxの座標がどれだけ離れているのか）が教師データとなる。

Bbox(xmin,ymin,xmax,ymax)

Dbox(cx,cy,w,h)

True(Δcx , Δcy , Δw , Δh)

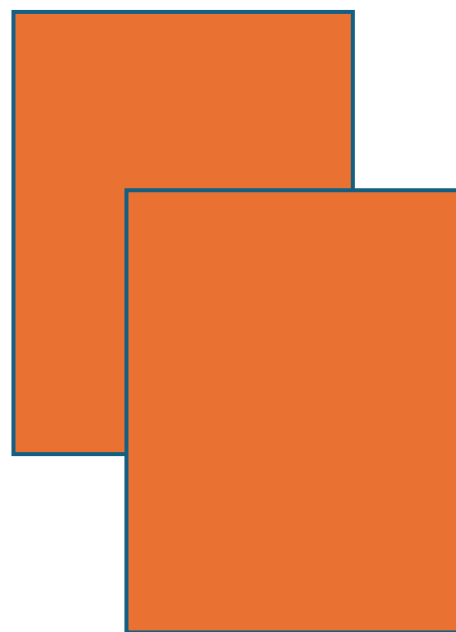


IoU(Intersection over Union)値とは

IoU値 = $(A \cap B \text{の面積}) / (A \cup B \text{の面積})$



$A \cap B$



$A \cup B$

学習と推論の流れ

・推論

- ①検出したいデータの前処理をする。
- ②前処理したデータをネットワークに入れる。各Dboxごとにlocとconfを求める。
- ③ locをもとにDboxをBboxに変換する。その後、各クラスごとにIoU値を利用して同じ物体を指しているか調べる。同じ物体を指している場合は、最も大きいconfを持つBbox以外は削除する。つまり、各物体はそれぞれのクラスの最もconfが大きいBboxを持つ。
- ④残ったBboxの座標をもとに、元のデータ上にBboxを描写する。

実際にVOC2012を使って、学習と推論

- VOC2012について

クラス数は20（人、犬、猫など）

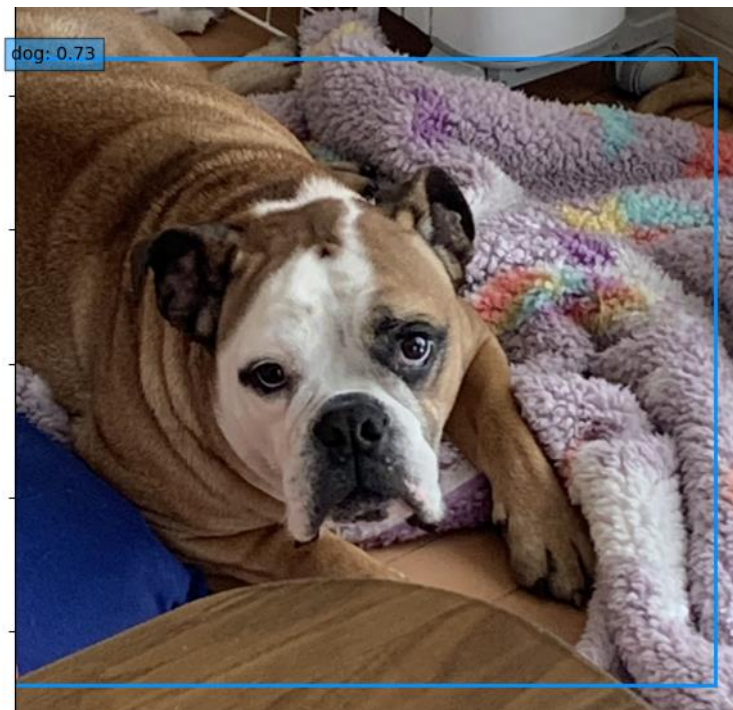
データの数約11,540枚

- 時間の都合上、30epochで切り上げた。まだ、過学習は起こっていなかったため、epoch数を増やすことで精度はもっとあげられた。

（過学習とは、訓練画像に過剰に適用しすぎて、汎化性能を失ってしまうこと）

結果

- 以下のように検出に成功した。



オリジナル画像で物体検出

公開されている画像だけではなく、自分で写真を撮り、バウンディングボックスを作って物体検出をした。

今回作ったもの

「お気に入りのグミを検出するai」

練習で作ったものなので、学習に使用したデータ数は少ないが、検出する種類は少ないため精度はそこそこ。ただし、学習データに背面を映した写真は入れてないので検出は表面に限る。

学習用データ作り

お気に入りのグミ

“つぶぐみ”、“かむかむレモン（グミ？）”

この2種類を検出できるようにしたい。

それぞれ350枚ずつ（150枚は一つ入り、残り200枚は二つ入る）

学習データ数:評価データ数 = 8:2

使用したアノテーションツール Vott

結果

以下のように検出できた。



結果（2）

しかし、以下のようにうまく検出できないこともあった。



次はしっかり時間をかけて、検出する種類が多く、高い精度を持つものを作りたい。