

距離学習を導入した CenterNet による 腹部超音波画像からの肝腫瘍検出と分類

B4 原 英吾

1 研究背景および目的

● 背景

- 器具の操作と診断を同時に行わなければならない高難易度
- 肝臓は沈黙の臓器と呼ばれ初期には自覚症状がほとんどない
 - * 症状を自覚している時には重症化している場合が多い
- 機械学習による診断のサポート
 - * 良性・悪性を見分けることが重要視される
 - * 図 1 の様に明らかなラベル不足¹のある画像が存在する

● 目的

- 既存の研究を踏まえたモデルの精度向上
 - * 良性・悪性判別の高精度化
- 超音波支援システムの開発
 - * 早期発見につながると良い

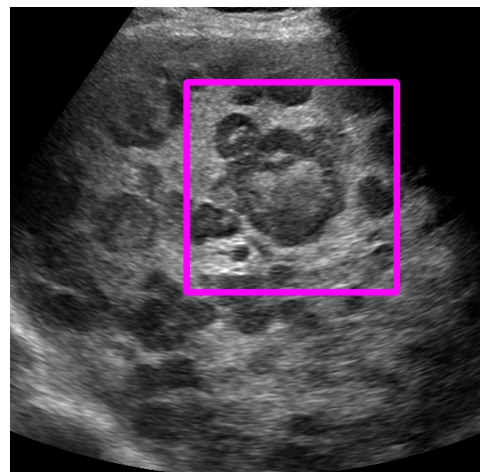
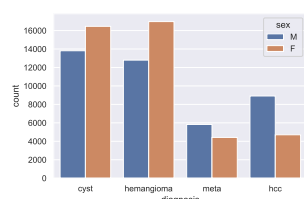


図 1: ラベル不足のある診断画像例

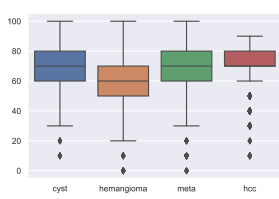
2 これまでの研究のまとめ

● データセット

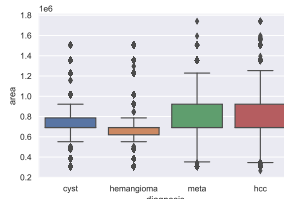
- 国立研究開発法人日本医療研究開発機構 (AMED)²が提供している延べ 8 万枚に及ぶ以下のデータが付随
 - * 腹部超音波画像, ROI
 - * 年齢, 性別



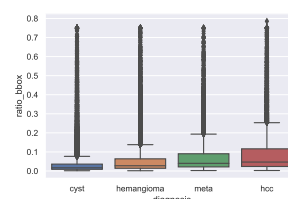
(a) 性別毎の画像枚数



(b) 診断名毎の年齢分布



(c) 診断名毎の画像サイズの分布



(d) 診断名毎の bbox の割合

図 2: データセットにおけるデータの分布

— 性別 (図 2a)

- * HCC(肝細胞癌) は男性が罹患しやすい
 - ・ 昔は男性の方が飲酒・タバコが多く癌に罹りやすかったという時代背景があるかもしれない
- * hemangioma(血管腫) は女性が罹患しやすい
- * Meta(転移性肝癌) は他の症状よりも少ない

— 年齢 (図 2b)

- * cyst(単純嚢胞), hemangioma(血管腫) の分布にはあまり特徴がない

¹今回は図 1 の様なアノテーションが不足しているものを指す

²<https://www.amed.go.jp/>

- * hemangioma(血管腫) は比較的若年層でも罹患する
- * Meta(転移性肝癌) における 0 歳はラベルミスである可能性が高い
- * HCC(肝細胞癌) は比較的高齢者が罹患しやすい
- 画像サイズ (図 2c)
 - * hemangioma(血管腫) は比較的画像サイズが統一されている
 - ・ 腫瘍の大きさが血管に依存するためあまり偏りが生じていない?
- bbox の画像に占める割合 (図 2d)
 - * cyst(単純嚢胞) は他の診断と比べて bbox の割合が低い ($\frac{1}{2}$ 程度) である
 - * HCC(肝細胞癌) は画像に占める bbox の割合が高い
- 症状毎の特徴を調査

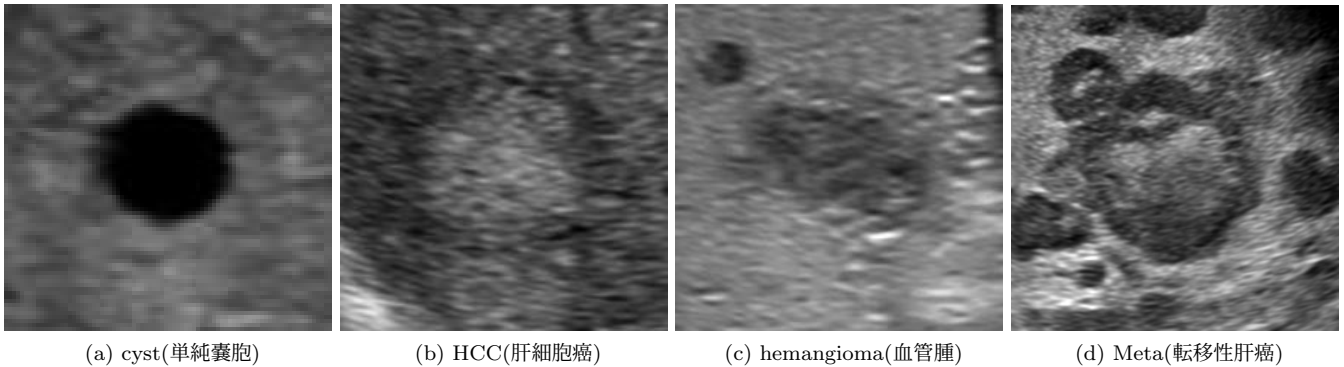


図 3: 症状毎における腫瘍の超音波画像

- cyst(単純嚢胞) (図 3a)
 - * 液体が貯留されている状態の良性腫瘍
 - * 症状がでないことが多いため大きな腫瘍になって発見されることが多い
 - * 嚢胞の内腔に向けて増殖するため転移することは少ない
- HCC(肝細胞癌) (図 3b)
 - * 肝臓にできる**悪性腫瘍**の中で最も多いと言われている
 - * 約 90%がウイルス感染症が原因
 - ・ B 型肝炎ウイルス (HBV) が約 20%
 - ・ C 型肝炎ウイルス (HCV) が約 70%
- hemangioma(血管腫) (図 3c)
 - * 肝臓にできる良性腫瘍の中で最も多い
 - * 女性ホルモンが原因で女性が罹患しやすいと言われているが詳しくは解明されていない
 - * 血管が無数に絡み合うことによって出来た血管の塊であることから血流が遅いという特徴がある
 - * 他の臓器に浸潤したり転移することは無いと言われている
- Meta(転移性肝癌) (図 3d)
 - * 門脈を介して大腸癌などの消化器癌から転移する割合が多い**悪性腫瘍**
 - * 類似したエコーパターンをもつ腫瘍が多発してみられることが多い
- データクレンジング
 1. 400 × 400 以下の画像の除外
 2. Perceptual Hash を利用した類似画像の除外
 3. 青色や黄色のスケールの除去
- 提供されているデータを COCODataset の形式に変換
 - train data : test data : val data = 67122 : 8390 : 8391
 - 見やすいようにインデントしたファイル³も作成

³[//aka/work/hara.e/AMED/lib/dataset/annotations/train_large.json](https://aka/work/hara.e/AMED/lib/dataset/annotations/train_large.json) など

- 1 クラスと 4 クラス検出の学習結果

- 評価指標の値を算出する為の条件

- * IoU の閾値は 0.25
 - * 信頼度は 0.00 から 1.00 まで 0.05 区切りで F1 値が最大となる値を指定
 - ・ YOLOX: 0.35
 - ・ CenterNet: 0.30
 - * 腫瘍に複数の予測がかかった時
 - ・ NMS で統合されていなくても正検出を 1 と扱う
 - ・ 異なるクラスが予測された場合は信頼度が最大となる予測を検出として扱う

| | | | | | | |
|-------|------|-------|------|-----|------|------|
| Label | None | 0 | 851 | 371 | 468 | 28 |
| | cyst | 297 | 2708 | 17 | 15 | 0 |
| | HCC | 371 | 25 | 810 | 152 | 13 |
| | hem. | 591 | 33 | 126 | 2203 | 9 |
| | Meta | 339 | 37 | 396 | 141 | 108 |
| | | None | cyst | HCC | hem. | Meta |
| | | Preds | | | | |

(a) YOLOX[2] での 4 クラス検出結果

| | | | | | | |
|-------|------|-------|------|-----|------|------|
| Label | None | 0 | 619 | 147 | 398 | 173 |
| | cyst | 208 | 2791 | 6 | 18 | 14 |
| | HCC | 271 | 12 | 871 | 103 | 114 |
| | hem. | 269 | 20 | 49 | 2558 | 66 |
| | Meta | 265 | 12 | 115 | 80 | 549 |
| | | None | cyst | HCC | hem. | Meta |
| | | Preds | | | | |

(b) CenterNet[4] での 4 クラス検出結果

図 4: YOLOX[2] と CenterNet[4] での 4 クラス検出した時の混同行列

表 1: 1 クラス検出と 4 クラス検出の学習で得られた精度

| model | backbone | classes | epoch | size | batch_size | IoU | | | area | | |
|--------------|----------|---------|-------|------|------------|-------|------------------|------------------|-----------------|-----------------|-----------------|
| | | | | | | mAP | AP ₅₀ | AP ₇₅ | AP _S | AP _M | AP _L |
| YOLOX[2] | DarkNet | 1 | 300 | 512 | 64 | 0.519 | 0.839 | 0.558 | - | 0.639 | 0.631 |
| | | 4 | | | | 0.279 | 0.526 | 0.248 | - | 0.221 | 0.288 |
| CenterNet[4] | ResNet18 | 4 | 300 | 512 | 16 | 0.344 | 0.639 | 0.332 | - | 0.347 | 0.326 |

表 2: YOLOX[2] で 4 クラス検出を行った時の評価指標毎の値

| Diagnosis | データ総数 | 過検出数 | 未検出数 | precision | recall | f1-score |
|------------|-------|------|------|-----------|--------|----------|
| cyst | 3037 | 851 | 297 | 0.7411 | 0.8917 | 0.8094 |
| HCC | 1371 | 371 | 371 | 0.4709 | 0.5908 | 0.5241 |
| hemangioma | 2962 | 468 | 591 | 0.7395 | 0.7438 | 0.7416 |
| Meta | 1021 | 28 | 339 | 0.6855 | 0.1058 | 0.1832 |
| 合計 | 8391 | 1718 | 1598 | 0.6588 | 0.5830 | 0.5646 |

表 3: CenterNet[4] で 4 クラス検出を行った時の評価指標毎の値

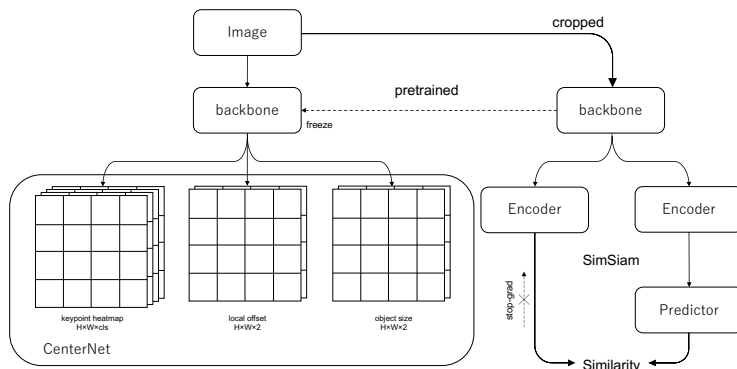
| Diagnosis | データ総数 | 過検出数 | 未検出数 | precision | recall | f1-score |
|------------|-------|------|------|-----------|--------|----------|
| cyst | 3037 | 619 | 208 | 0.8080 | 0.9190 | 0.8600 |
| HCC | 1371 | 147 | 271 | 0.7332 | 0.6353 | 0.6807 |
| hemangioma | 398 | 269 | 235 | 0.8103 | 0.8636 | 0.8361 |
| Meta | 1021 | 173 | 265 | 0.6339 | 0.5377 | 0.5669 |
| 合計 | 8391 | 1208 | 979 | 0.7377 | 0.7389 | 0.7359 |

- 実験で用いる SimSiam[5] のプログラムを書いた

- 元論文のプログラムでは backbone の fc 層しか用いられていなかった
 - (自分の書いたプログラムではないからかもしれないが) 汚くて使いにくかった

3 前回の LT からの進捗

- backbone を SimSiam[5] で距離学習後に重みを全て固定して CenterNet[4] で 4 クラス検出を行った



(a) 実験で使用する SimSiam[5] と CenterNet[4] の複合モデル概要

| | | | | | |
|-------|------|------|-----|------|------|
| Label | None | cyst | HCC | hem. | Meta |
| None | 0 | 1871 | 398 | 725 | 229 |
| cyst | 327 | 2649 | 14 | 29 | 18 |
| HCC | 804 | 33 | 336 | 125 | 73 |
| hem. | 1574 | 75 | 93 | 1168 | 52 |
| Meta | 579 | 51 | 115 | 90 | 186 |
| | None | cyst | HCC | hem. | Meta |

(b) backbone の重みを全て固定したモデルでの混同行列

図 5: 実験で使ったモデルとその検出と分類の結果

- 図 5b から (図 4b と比べて)

- * backbone の重みを固定した影響で過検出・未検出数が上昇していることが明らか
- * 分類精度も上昇していないので特徴量抽出層 (backbone) の重みを固定した効果が得られていない

表 4: SimSiam[5] で距離学習後に重みを固定して CenterNet[4] で 4 クラス検出を行った時の評価指標毎の値

| Diagnosis | データ総数 | 過検出数 | 未検出数 | precision | recall | f1-score |
|------------|-------|------|------|-----------|--------|----------|
| cyst | 3037 | 1871 | 327 | 0.5661 | 0.8722 | 0.6866 |
| HCC | 1371 | 398 | 804 | 0.3513 | 0.2451 | 0.2888 |
| hemangioma | 398 | 725 | 1574 | 0.5466 | 0.3943 | 0.4581 |
| Meta | 1021 | 229 | 579 | 0.3333 | 0.1822 | 0.2356 |
| 合計 | 8391 | 3223 | 3284 | 0.4494 | 0.4235 | 0.4173 |

- 表 4 から (表 3 と比べて)

- * 全体的に 30% 程精度が低くなっている
- * 特に HCC は 40% 程精度が低くなっている

- 考察

- * 事前に対照学習をした効果が十分に得られていないので下流層から上流層に向けて段階的に重みの固定を解除していくのが良さそうなのではないか
 - ・ 現在学習途中
 - ・ 決まった epoch 数学習するのではなく mAP の上昇が止まったところで次層の固定を解除することも試してみたい
- * 検出と分類の head を分けて実験したらどうなるのか気になる
 - ・ 固定した重みが検出と分類を同時にするには有効ではない可能性があったから
 - ・ 今後のためにも CenterNet[4] のプログラムを実装していく

- CenterNet[4]⁴のプログラムを実装

- model のプログラムは完成
- loss のプログラムは途中

⁴<https://github.com/a5chin/centernet>

- COCOAPI (validation データ) での結果まとめ

表 5: 1 クラス検出と 4 クラス検出の学習で得られた精度

| model | backbone | weights | classes | epoch | size | IoU | | | area | | |
|--------------|----------|---------|---------|-------|------|-------|------------------|------------------|-----------------|-----------------|-----------------|
| | | | | | | mAP | AP ₅₀ | AP ₇₅ | AP _S | AP _M | AP _L |
| YOLOX[2] | DarkNet | flex | 1 | 300 | 512 | 0.519 | 0.839 | 0.558 | - | 0.639 | 0.631 |
| | | | 4 | | | 0.279 | 0.526 | 0.248 | - | 0.221 | 0.288 |
| CenterNet[4] | ResNet18 | flex | 4 | 300 | 512 | 0.344 | 0.639 | 0.332 | - | 0.347 | 0.326 |
| | | fix | | | | 0.150 | 0.309 | 0.127 | - | 0.119 | 0.161 |

4 今後の課題&スケジュール

- 7/12 までに
 - 対照学習をさせた backbone の重みを徐々に解除していきながら CenterNet[4] を学習させる
 - * epoch 数を決めて (50epoch ずつ)
 - * 解除した layer の重みを学習しきれてから次の層の学習に移る
 - できれば CenterNet[4] の loss を実装する

参考文献

- [1] Anonymous authors. [DEEP DOUBLE DESCENT: WHERE BIGGER MODELS AND MORE DATA HURT](#) ICLR 2020, 2020.
- [2] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun. [YOLOX: Exceeding YOLO Series in 2021](#), 2021.
- [3] Ting Chen, Simon Kornblith, Mohammad Norouzi, Geoffrey Hinton. [A Simple Framework for Contrastive Learning of Visual Representations](#), 2020.
- [4] Xingyi Zhou, Dequan Wang, Philipp Krähenbühl. [Objects as Points](#), 2019.
- [5] Xinlei Chen, Kaiming He. [Exploring Simple Siamese Representation Learning](#), 2020.