

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
 - 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
 - 先行研究では 0.7617 ($\approx 0.8550 \times 0.8909$)
 - * 検出 (Recall) : 0.8550
 - * 分類 (Accuracy) : 0.8909
- 目的
 - 既存の研究を踏まえたモデルの精度向上
 - 超音波支援システムの開発

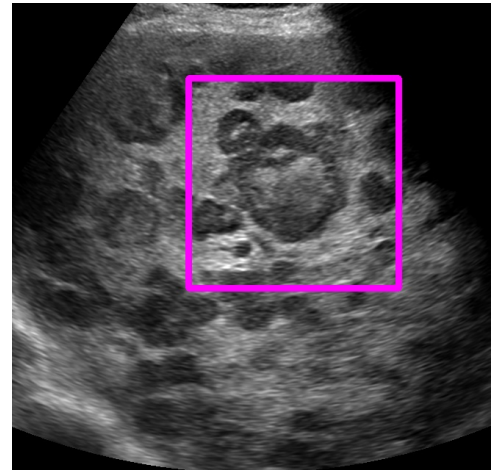


図 1: ラベル不足のある診断画像例

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

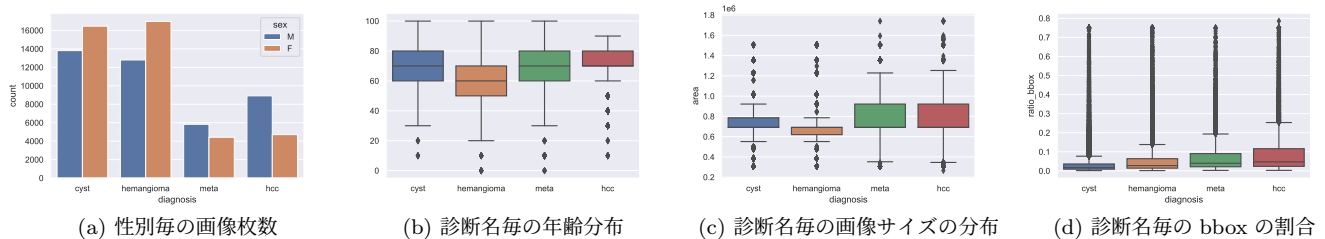


図 2: データセットにおけるデータの分布

- 性別 (図 2(a))
 - * 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 2(b))
 - * 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 2(c))
 - * 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 2(d))
 - * 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

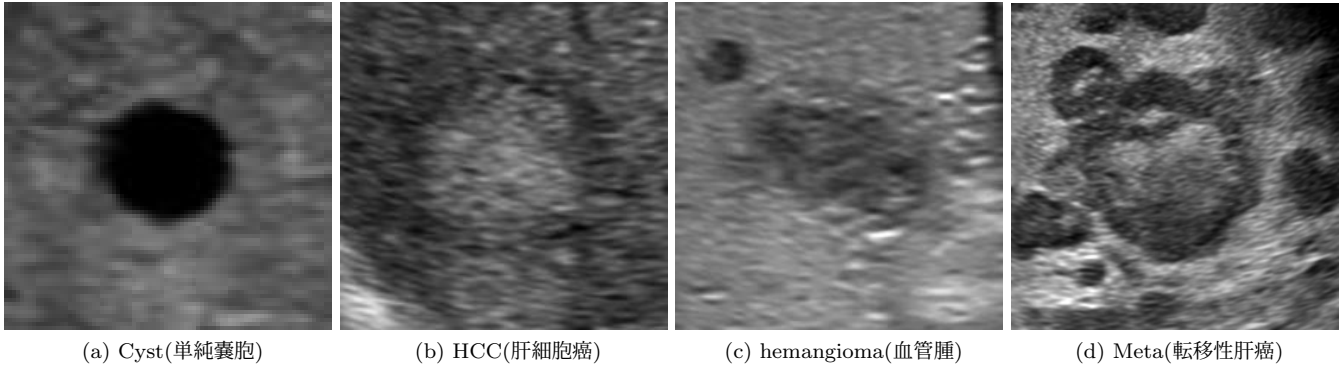


図 3: 症状毎における腫瘍の超音波画像

- Cyst(単純嚢胞) (図 3(a))
 - * 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 3(b))
 - * 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 3(c))
 - * 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 3(d))
 - * 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

3 前回の GM からの進捗

- モデル・クラス数ごとの検出結果の比較

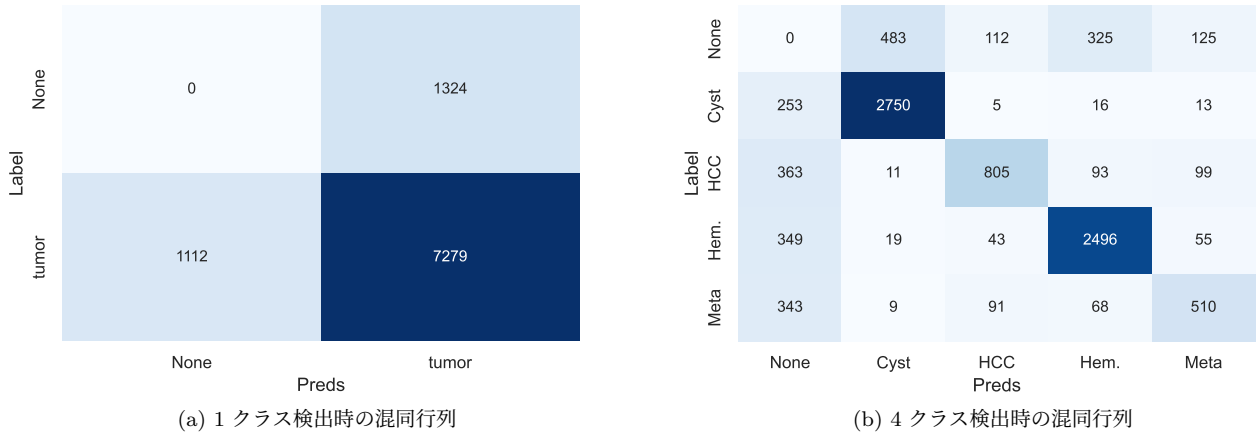


図 4: CenterNet[3] で検出した時の混同行列

表 1: SimSiam[2] で距離学習後に段階的に backbone の重みを解除したモデルでの評価指標毎の値

Diagnosis	データ総数	FP	FN	precision	recall	f1-score
Cyst	3037	483	253	0.8405	0.9055	0.8718
HCC	1371	112	363	0.7623	0.5872	0.6634
Hemangioma	398	325	349	0.8326	0.8427	0.8376
Meta	1021	125	343	0.6359	0.4955	0.5595
合計	8391	1045	1308	0.8072	0.7819	0.7944

表 2: 実験条件と test データでの Recall

model	classes	IoU	confidence	Recall
YOLOX[1]	1	0.25	0.45	0.9045
	4		0.40	0.6691
CenterNet[3]	1	0.25	0.35	0.8675
	4		0.35	0.7819

表 3: COCO API での評価

model	backbone	classes	epoch	size	batch_size	IoU			area ⁴		
						mAP	AP ₅₀	AP ₇₅	AP _S	AP _M	AP _L
YOLOX[1]	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[3]	ResNet18	1	300	512	16	0.396	0.750	0.366	-	0.419	0.389
		4				0.344	0.639	0.332	-	0.347	0.326

● 考察

- 1 クラス検出の時に YOLOX[1] の精度の方が高い（表 2, 表 3）のは Double Descent[4] が起きているから？
 - * Noisy Label が主な原因
- 1 クラス検出の方が 4 クラス検出に比べて精度が高い
 - * 4 クラス検出には 1 クラス検出に加えて分類問題が入るので当たり前

⁴ $S < 32 \times 32 \text{ pix}^2 < M < 96 \times 96 \text{ pix}^2 < L$

4 今後の課題&スケジュール

- 10/25 までに
 - WiNF2022 の申し込み締切
- 12/1 までに
 - IWAIT2023 の full-paper 提出期限

付録 A CenterNet

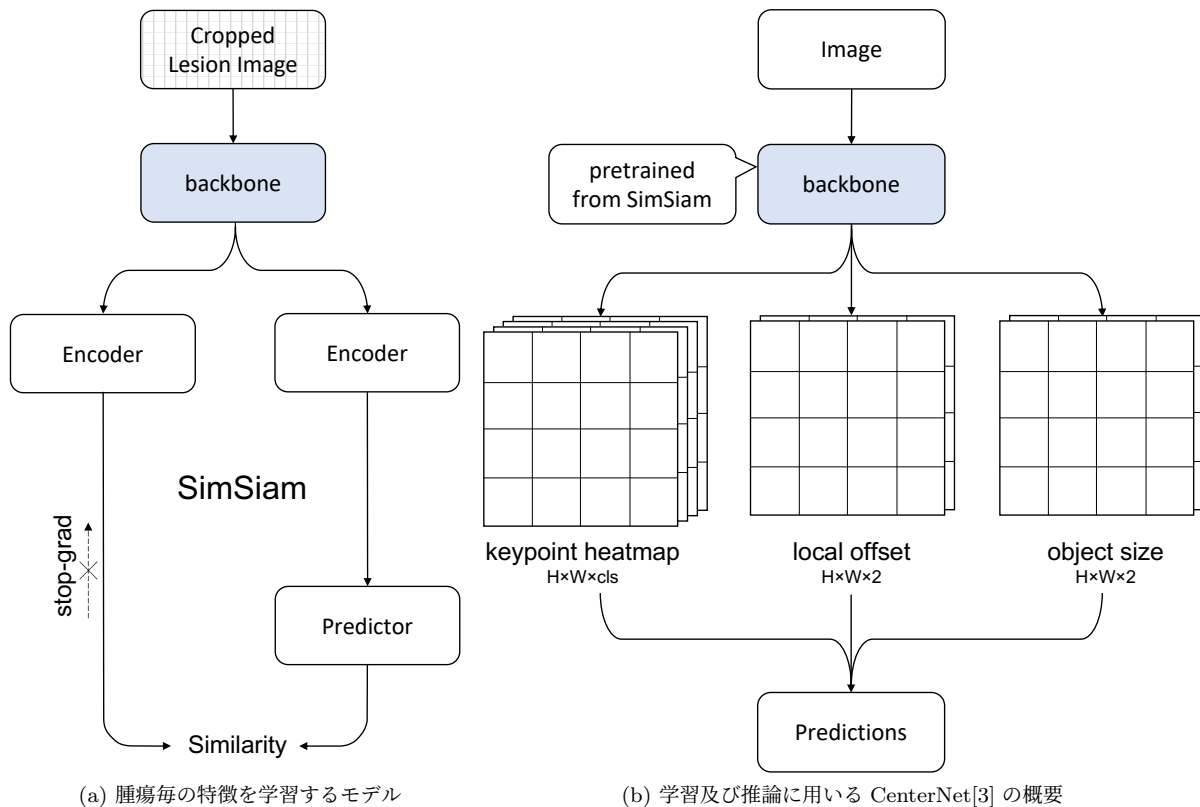


図 5: 4 クラス検出の実験で使ったモデル

参考文献

- [1] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun: “YOLOX: Exceeding YOLO Series in 2021,” arXiv preprint, arXiv:2107.08430, 2021.
- [2] X. Chen and K. He: “Exploring simple siamese representation learning,” Proc. 2021 IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 15750–15758, 2021.
- [3] X. Zhou et al.: “Objects as Points,” arXiv preprint, arXiv:1904.07850, 2019.
- [4] Nakkiran, Preetum and Kaplun, Gal and Bansal, Yamini and Yang, Tristan and Barak, Boaz and Sutskever, Ilya: “Deep double descent: Where bigger models and more data hurt” Journal of Statistical Mechanics: Theory and Experiment, 124003, 2020.