

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

原 英吾[†] 道満 恵介[†] 目加田 慶人[†] 西田 直生志[‡] 工藤 正俊[‡]

[†] 中京大学 工学部

Email: hara.e@md.sist.chukyo-u.ac.jp, {kdoman,y-mekada}@sist.chukyo-u.ac.jp

[‡] 近畿大学 医学部

Email: {naoshi,m-kudo}@med.kindai.ac.jp

1 はじめに

2020 年には、世界中で 83 万人以上の方が肝臓がんで死亡しており、世界のがん死亡原因の第 3 位である。また、近年、超音波検査は肝がんの初期診断に有効な方法と考えられている。しかし、器具の操作と診断を同時に行わなければならないことから腹部超音波検査は難易度が高いとされている。そのため、腹部腫瘍の検出や分類などはコンピュータを用いた診断が求められている。

従来は、図 1(a) の様な与えられた超音波画像に対して腫瘍の検出後に、切り出された腫瘍領域 [1] を分類する [2] という 2 段階の推論ステップによって構成されている。

山岸らの先行研究 [1] により、YOLOv5 による腫瘍検出手法が提案されている。YOLOv5 とは、Ultralytics 社により開発されており、2020 年 6 月に公開された画像中の物体認識ライブラリである。この先行研究では、IoU しきい値が 0.25 で信頼度のしきい値を 0.3 とした時に、再現率、適合率、F1 スコアがそれぞれ 0.855, 0.896, 0.875 となることが報告されている。しかし、YOLO 系統のモデルは様々なアスペクト比の物体を検出するという前提のため、本稿の問題設定に適していない可能性が高い。

西田らの先行研究 [2] により、図 1(b),1(c),1(d),1(e) の様な切り取られた腫瘍画像に対して、VGG16[3] ベースの CNN モデルによる腫瘍分類手法が提案されており、accuracy が 0.8909 であることが報告されている。ここで、検出段階において画像を切り取らないことで、分類時に大域特徴を使用できるため精度の向上が見込めると考える。

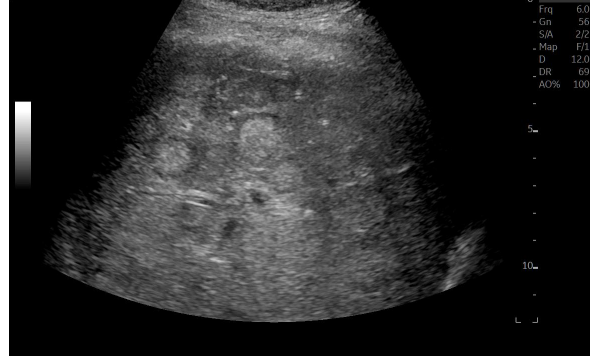
本稿では、図 1 に示すように、単純嚢胞 (図 1(b)), 血管腫 (図 1(c)), 肝細胞がん (図 1(d)), そして転移性肝がん (図 1(e)) の検出と分類を同時に行う。なお、後者 2 症状は悪性であるため極力未検出を抑制する必要がある。

2 提案手法

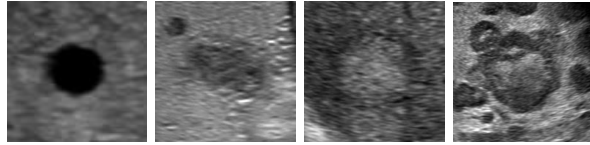
本稿では、切り取られた腫瘍画像を入力として SimSiam[4] で距離学習をさせたバックボーンネットワークを用いて CenterNet[5] をファインチューニングし、1 段階の推論手法を提案する。図 2 に示すように、腹部超音波画像に影像される肝腫瘍は形状が円形であることから、CenterNet のみを用いて推論することで効果的に腫瘍の検出・分類できると考える。

CenterNet に用いるバックボーンネットワークは、切り取られた腫瘍の画像を入力として SimSiam を用いた距離学習を行う。その後、1(a) の様に全体が映っている画像を用いてファインチューニングする。SimSiam は単純な構造を持ち、少ないパッチサイズでも検出と分類の両タスクに有効な特徴を学習することが可能な上、図 1 に示す様に 4 種類の腫瘍が画像の見た目が大きく異なっていることから採用した。

また、従来手法よりも推論に使用するモデルを減らすことで、メモリの削減や推論速度の上昇が見込める上に、超音波画像全体を



(a) 実際の診断で使われる画像



(b) 単純嚢胞 (c) 血管腫 (d) 肝細胞がん (e) 転移性肝がん

図 1: 実験で用いる超音波画像例

表 1: データセット分割後のそれぞれの枚数

診断名	train	validation	test	合計
単純嚢胞	24,217	3,018	3,037	30,272
血管腫	23,749	3,059	2,962	29,770
肝細胞がん	10,934	1,310	1,371	13,615
転移性肝がん	8,222	1,003	1,021	10,246
合計	67,122	8,390	8,391	83,903

用いて腫瘍の分類も行うため 4 クラス検出の高精度化が見込める。

このような枠組みにより、本手法は入力画像から肝腫瘍の位置や大きさ、種類 (keypoint heatmap) を出力している。

3 実験

本節では、3.1 節で作成したデータセットを用いて、3.3 節の学習を行う。そして、最後の 3.4 節で提案手法の有効性を確かめる評価実験を行った。

3.1 データセット

本実験では超音波医学会 (AMED) から提供されている全 89,903 枚の腹部超音波画像を、表 1 で示す様に train : test : validation = 8 : 1 : 1 と分割してモデルを学習させ、評価を行った。

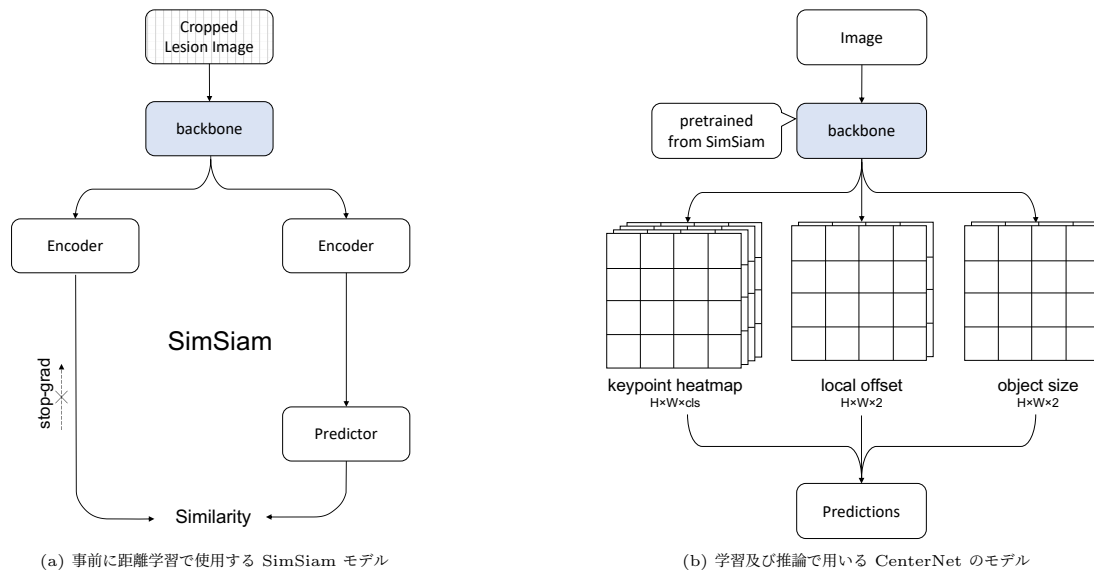


図 2: 提案手法の構造: バックボーンネットワークの事前学習に用いる SimSiam と推論を行う CenterNet

表 2: 学習で使用するモデルとそのハイパーパラメータ

	SimSiam	CenterNet
バッチサイズ	64	16
エポック	300	300
学習率	1.0×10^{-4}	5.0×10^{-5}
入力サイズ	512×512	512×512
最適化関数	SGD	SGD
誤差関数	Cosine Similarity	L1 Gaussian Focal

表 3: 提案手法を用いたモデルでの評価指標毎の値

診断名	腫瘍数	FP	FN	適合率	再現率	F1
Cyst	3,037	492	357	0.832	0.873	0.852
HCC	1,371	168	339	0.722	0.601	0.657
Hem.	398	301	465	0.828	0.796	0.812
Meta	1,021	172	290	0.619	0.543	0.578
合計	8,391	1,133	1,451	0.750	0.703	0.725

表 4: YOLOX での評価指標毎の値

診断名	腫瘍数	FP	FN	適合率	再現率	F1
Cyst	3,037	851	297	0.741	0.892	0.809
HCC	1,371	371	371	0.471	0.591	0.524
Hem.	2,962	468	591	0.740	0.744	0.742
Meta	1,021	28	339	0.686	0.106	0.183
合計	8,391	1,718	1,598	0.685	0.695	0.690

3.2 データセットの問題

このデータセットの特性として、以下の 3 点が挙げられる。

- (1) 画像内にカラーのスケールが映り込んでいる場合がある
 - (2) ほぼ同じ画像が混ざっている
 - (3) 複数の腫瘍が写っていてもアノテーションが 1 つしかない
- (1), (2) については頑健性低下の大きな要因になるため、それぞれ中島ら [6] と同様の手法でデータセットを作成した。具体的には、スケールを乱数で消すことや Hash 値による類似画像の削除により、データクレンジングを行っている。また、(3) に関しては、図 1(e) に示す様に、転移性肝がんの他部位にも同様に転移しやすいという性質により、主に転移性肝がんの検出精度を下げる大きな原因になることがわかっている。

3.3 提案モデルの学習

2 節で提案した SimSiam と CenterNet を 3.1 節で作成した学習データで学習させた。その際に、表 2 に示すハイパーパラメータを用いた。なお、CenterNet の誤差関数に関しては、keypoint heatmap で Gaussian Focal Loss を、local offset と object size で L1 損失を用いているため複数記述している。

また、SimSiam で学習させた初期重みを有効活用させるために、CenterNet の学習ではバックボーンネットワークの重みを固定し、学習を進めるにつれて段階的に解除した。

3.4 評価実験

学習が終了したモデルについては、表 1 に示されたテストデータを用いて 4 クラス検出の評価実験を行った。その際、先行研究に則り、IoU のしきい値は 0.25、信頼度のしきい値は 0.00 から 1.00 まで 0.05 ずつ変化させて計測した F1 値が最大となる時の値とする。今回の評価実験では、信頼度のしきい値は 0.40 となった。

4 結果

3.4 節の条件のもとでの検出結果 (図 3)、提案手法の混同行列 (図 4)、そしてその集計データ (表 3) を示す。なお、図 3 と図 4 に関しては、単純嚢胞 (Cyst)、肝細胞がん (HCC)、血管腫 (Hem.)、転移性肝がん (Meta) としている。また、比較手法として YOLOX[7] で 4 クラス分類を行った結果を図 4(b)、表 4 に示す。そして、表 5 に、手法毎における使用メモリ、推論速度、そして 4 クラス検出精度の比較結果を示す。

表 5: 従来手法, YOLOX と提案手法の比較

	使用メモリ (MB)	推論速度 (/sec)	再現率 $_{det}$	精度 $_{cls}$	4 クラス検出精度	信頼度
従来手法 [1, 2]	3898.3	0.0272	0.8550	0.8909	0.7617 ($\approx 0.8550 \times 0.8909$)	0.30
YOLOX[7]	4178.9	0.0167	0.8096	0.8447	0.6839 ($\approx 0.8096 \times 0.8447$)	0.35
提案手法	1838.2	0.0095	0.8271	0.9206	0.7614 ($\approx 0.8271 \times 0.9206$)	0.40

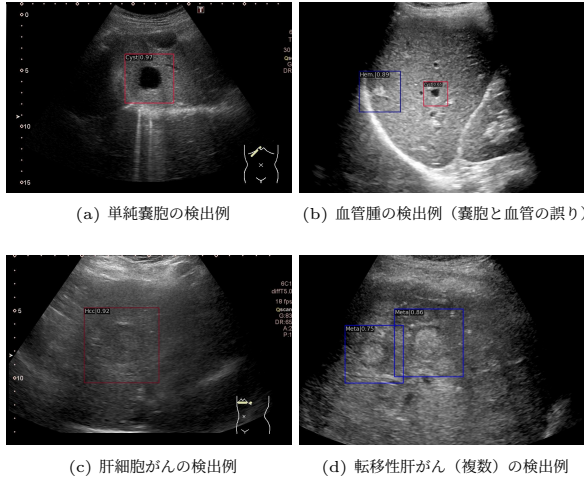


図 3: 提案手法での 4 クラス検出例

これらの結果から, 未検出, 過検出, 適合率, 再現率, 使用メモリ, そして推論速度の全ての評価項目において提案手法が YOLOX を上回っている。

5 考察

4 節から, 提案手法の再現率と Accuracy を算出し, 4 クラス検出結果の比較を行った結果を表 5 に示す。4 クラス分類精度が向上している要因として, SimSiam による距離学習や, 従来手法で得ることのできなかった画像の大域特徴を有効活用できたことが挙げられる。一方で, 検出精度が低下しているのは SimSiam によるバックボーンネットワークに対する 4 クラスの距離学習で検出に不向きな特徴量を学習してしまったのではないかと考えられる。また, 提案手法は従来手法よりも省メモリかつ高速に推論を行うことができることが示された。これにより, エッジデバイスでの実装が容易になったと考えられる。従って, 提案手法による一定の有効性を確認することができた。

表 3 に関して, 悪性腫瘍 (肝細胞がん, 転移性肝がん) の再現率が良性腫瘍 (単純嚢胞, 血管腫) に比べて低く, 診断が難しいと言える。これは, データセットに含まれる悪性腫瘍の数が少ない (表 1) ことによる学習データ不足が大きな原因だと考えられる。特に転移性肝がんの精度が低い問題に関しては, データの偏りに加えてアノテーション不足 (3.2 節) に依る所も大きいと推測する。

図 3 の様に, 高精度で 4 クラスの検出と分類ができていることが定性的にわかる。しかし, 血管を単純嚢胞 (Cyst) と誤検出 (図 3(b)) してしまったり, 未検出の腫瘍が複数存在する (図 3(d)) している。これらはそれぞれ, 図 1(b) に示す様に症状の特性が血管と似ていることや, 3.1 節で述べたアノテーションの問題が大きな要因であると考えられる。解決にはモデルの改善もそうだ

Label	None		492	168	301	172
	Cyst	357	2651	8	12	9
	HCC	339	8	825	97	102
	Hem.	465	23	57	2359	58
	Meta	290	13	84	80	554
		None	Cyst	HCC	Hem.	Meta
		Preds				

(a) 提案手法での混同行列

Label	None		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 Preds	Hem.	Meta

(b) YOLOX での混同行列

図 4: 提案手法と YOLOX で 4 クラス分類を行った際の比較

が, 実際の診断で用いるような超音波動画等の入力による腫瘍の追跡処理が必要であると考えられる。

6 まとめ

腹部超音波画像から肝腫瘍を同時に検出・分類する方法として, SimSiam に基づく距離学習を用いた CenterNet による 1 段階手法を提案した。この提案手法は, 従来の検出された腫瘍領域に対して 4 クラスの分類を行う, 2 段階手法と同等の高い分類精度を達成できることを確認した。また, この提案手法は従来手法よりも省メモリかつ高速に推論が可能である。

しかし, 分類精度は向上したが検出精度が低下してしまったことから, 腫瘍の検出と腫瘍の分類に必要な特徴量が異なっていると考えられる。したがって今後は, 検出と分類の両方に適したモデルとなるような改良をしていく予定であり, 例えば, 検出 head と分類 head をバックボーンネットワークから分離するなどの研究を行うことを想定している。

謝辞

本研究の一部は、AMED 革新的がん医療実用化研究事業および日本学術振興会科研費の援助による。

参考文献

- [1] Yamagishi, S., Doman, K., Mekada, Y., and Nishida, N., “Detection and tracking of liver tumors for ultrasound diagnostic support using deep learning,” Journal of Image and Graphics **10**(1), 50–55 (2022).
- [2] Nishida, N., Yamakawa, M., Shiina, T., Mekada, Y., Nishida, M., Sakamoto, N., Nishimura, T., Iijima, H., Hirai, T., Takahashi, K., et al., “Artificial intelligence (ai) models for the ultrasonographic diagnosis of liver tumors and comparison of diagnostic accuracies between ai and human experts,” Journal of gastroenterology **57**(4), 309–321 (2022).
- [3] Simonyan, K. and Zisserman, A., “Very deep convolutional networks for large-scale image recognition. cornell university (2014),” arXiv preprint arXiv:1409.1556 (2021).
- [4] Chen, X. and He, K., “Exploring simple siamese representation learning,” in [Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition], 15750–15758 (2021).
- [5] Zhou, X., Wang, D., and Krähenbühl, P., “Objects as points,” arXiv preprint arXiv:1904.07850 (2019).
- [6] Nakashima, T., Tsutsumi, I., Takami, H., Doman, K., Mekada, Y., Nishida, N., and Kudo, M., “A study on liver tumor detection from an ultrasound image using deep learning,” in [International Workshop on Advanced Imaging Technology (IWAIT) 2020], Lau, P. Y. and Shobri, M., eds., **11515**, 115151V, International Society for Optics and Photonics, SPIE (2020).
- [7] Ge, Z., Liu, S., Wang, F., Li, Z., and Sun, J., “Yolox: Exceeding yolo series in 2021,” arXiv preprint arXiv:2107.08430 (2021).