

論文輪講

U-Net: Convolutional Networks for Biomedical Image Segmentation

Olaf Ronneberger, Philipp Fischer, and Thomas Brox

知的高性能計算研究室
呂冰（ロビン）

概要

1. 研究背景

2. U-Net 紹介

2.1 データ 増強

2.2 処理の 流れ

2.3 U-Net の 特徴

2.4 計算

3. 実験

4. 実験結果

5. まとめ

研究背景

この論文で医療画像分割のために提案されたU-Netの特徴は

- 少数のデータでトレーニング可能
- 出力は認識結果ではなく、フィーチャマップ (feature map)
- 出力の2枚のフィーチャマップ(前景と背景)は元画像と同じサイズ

比較として、LeNetなど一般的な畳み込みネットワーク (CNN) は

- 大量のトレーニングデータが必要
- ネットワークの出力は単一のラベル (0~9)
- 抽出されたフィーチャマップは層数ごとに小さくなる

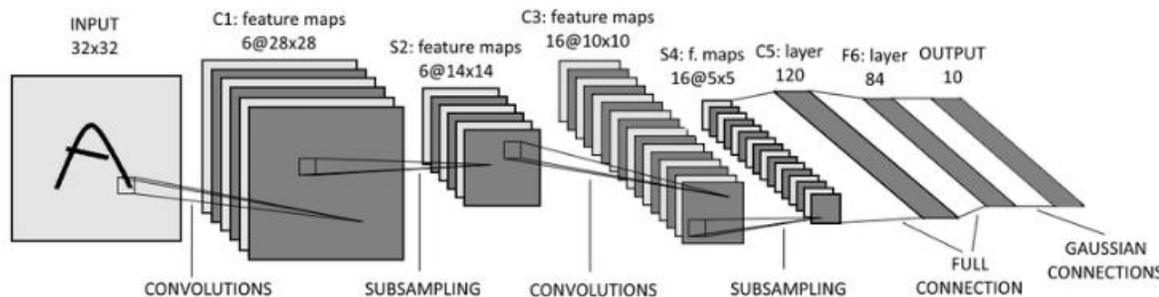


図1 LeNetの概要図

データ増強

なぜ少数のデータだけでトレーニングが可能なのですか？

➤ 弾性変形 (elastic deformation) によるデータ増強 (data augmentation)

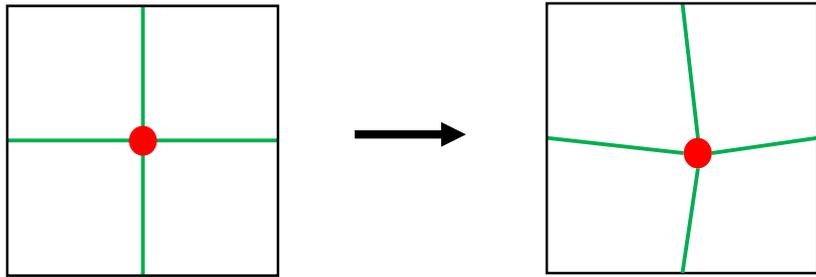


図1 弾性変形の概要

1. 図を均等に分割
2. 境界線の接点を一定の範囲内移動
3. 分割した各部分を変形させる

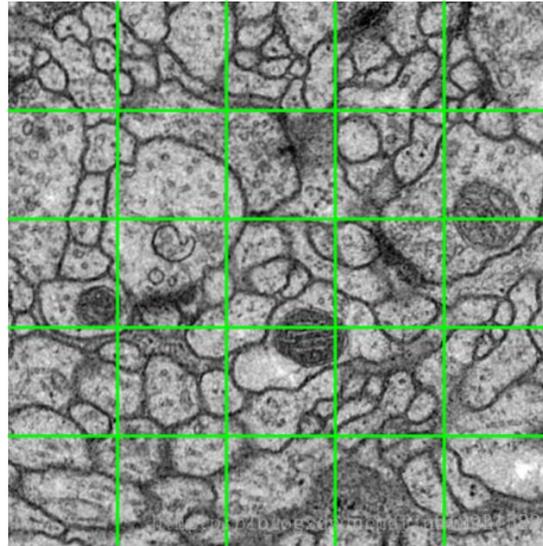


図2(a) 元画像

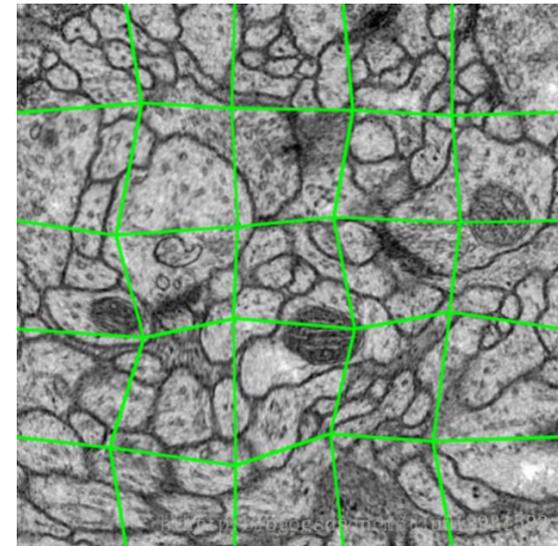
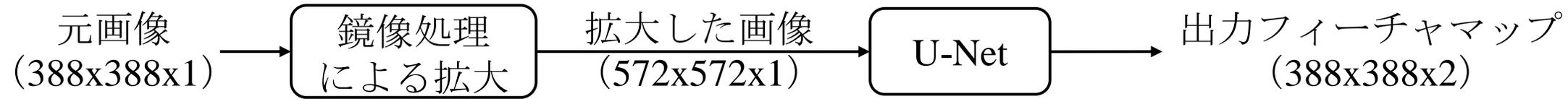


図2(b) 弾性変形した画像

ランダムな弾性変形により、1枚の元画像から複数枚の変形画像を生成し、
トレーニングデータを増強

処理の流れ



鏡像処理 (mirroring) による画像の拡大

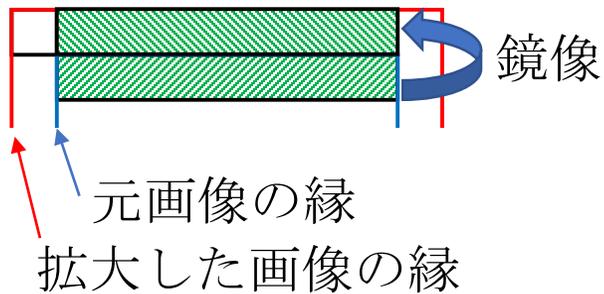


図1 鏡像処理の概要

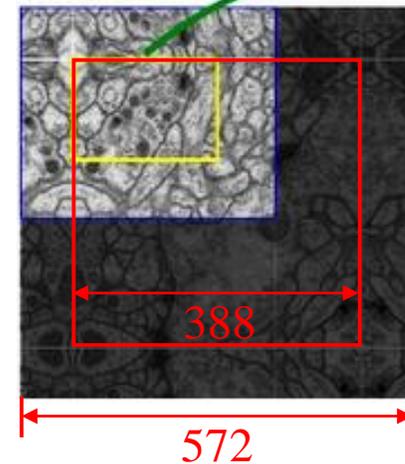


図2 拡大した画像

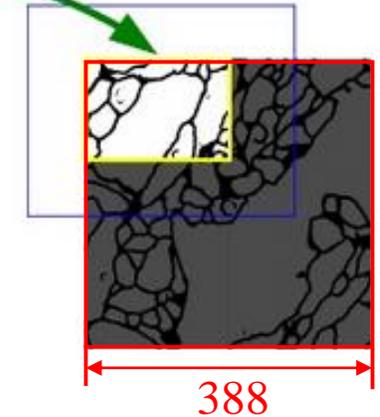


図3 出力
フィーチャマップ

パッチ (Patch) : 画像を分割 (図2、3青い枠)

❖ 黄色領域のセグメンテーションを予測するためには、青領域の入力が必要

U-Netの構成

1. 全層畳み込みネットワーク (Fully Convolution Network)
2. アップサンプリング (upsampling) と逆畳み込み (deconvolution) を用いて原画像サイズを得る。
3. 各ピクセルを分類

➤ 構成部分

- 特徴抽出部分：3x3畳み込み、ReLU、2x2マックスプーリング
- アップサンプリング部分：サンプリングを1回行うごとに、同じスケールのフィーチャマップと綴り合せ

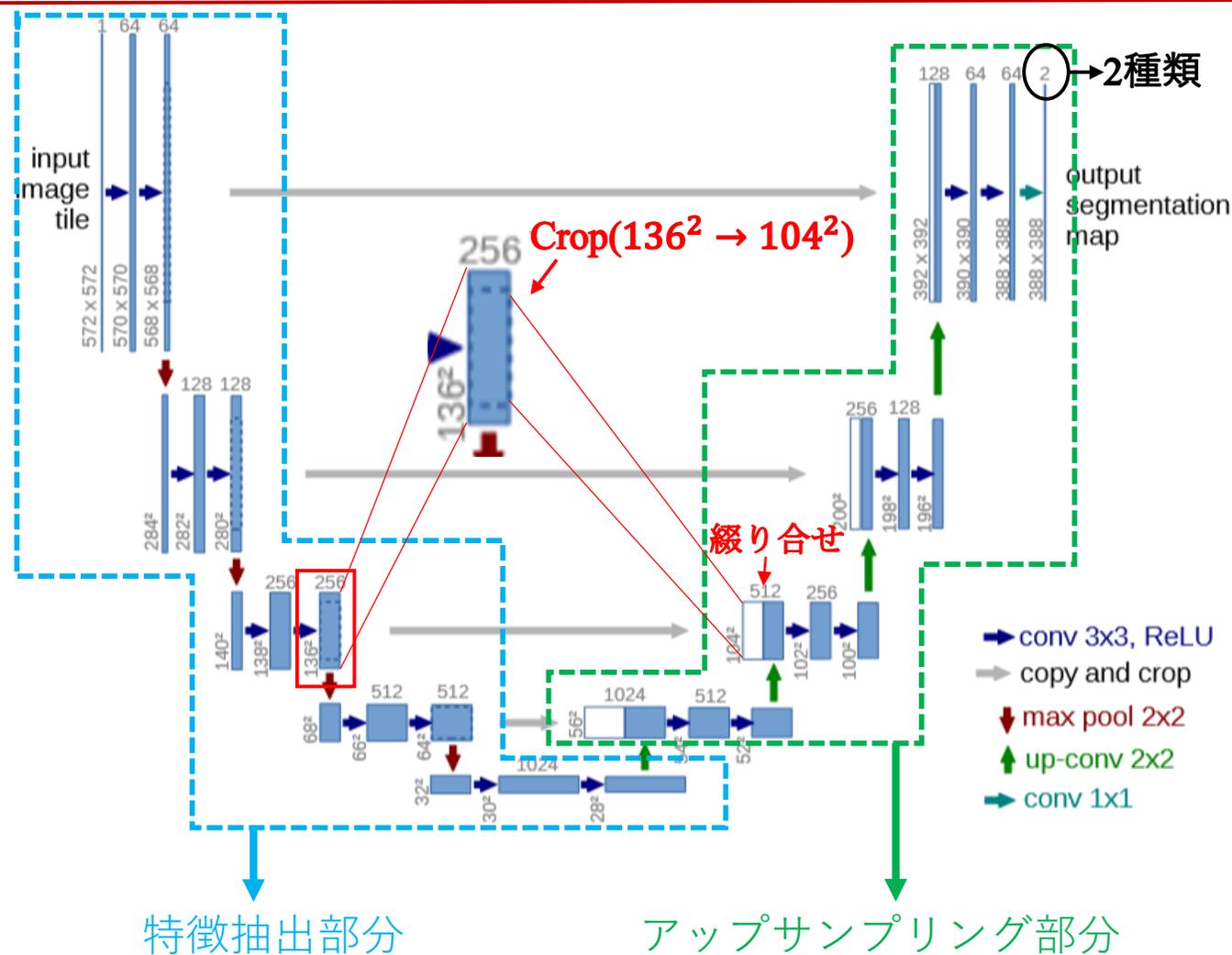


図1 U-Netの構成

計算

➤ 重みの計算

細胞境界付近のピクセルに大きい重みを与える

細胞境界から離れたピクセルに小さい重みを与える

$$w(\mathbf{x}) = w_c(\mathbf{x}) + w_0 \cdot \exp\left(-\frac{(d_1(\mathbf{x}) + d_2(\mathbf{x}))^2}{2\sigma^2}\right)$$

$d_1(x)$: ピクセルからその点に最も近い細胞境界までの距離

$d_2(x)$: このピクセルから2番目に近い細胞の距離

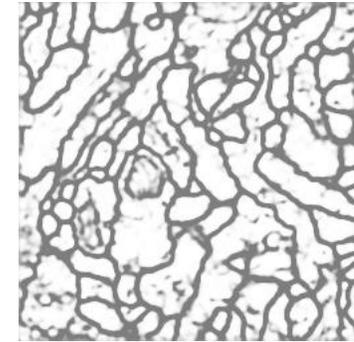


図1 U-Netの処理例
(黒い線は細胞の境界)

➤ ロス計算 (loss)

$$E = \sum_{\mathbf{x} \in \Omega} w(\mathbf{x}) \log(p_{l(\mathbf{x})}(\mathbf{x}))$$

x : あるピクセル

$l(x)$: このピクセルのラベル

$p_{l(x)}(x)$: このラベルにおける x の出力確率

論文の実験結果

- 電子顕微鏡記録におけるニューロン構造のセグメンテーション
前処理または後処理することなく0.0003529のワーピングエラーと0.0382のランドエラーを実現

Table 1. Ranking on the EM segmentation challenge [14] (march 6th, 2015), sorted by warping error.

Rank	Group name	Warping Error	Rand Error	Pixel Error
	** human values **	0.000005	0.0021	0.0010
1.	u-net	0.000353	0.0382	0.0611
2.	DIVE-SCI	0.000355	0.0305	0.0584
3.	IDSIA [1]	0.000420	0.0504	0.0613
4.	DIVE	0.000430	0.0545	0.0582
⋮				
10.	IDSIA-SCI	0.000653	0.0189	0.1027

- 光学顕微鏡画像における細胞セグメンテーション
77.5%の結果を達成
(46%の2番目にアルゴリズムより大幅に優れている)

Table 2. Segmentation results (IOU) on the ISBI cell tracking challenge 2015.

Name	PhC-U373	DIC-HeLa
IMCB-SG (2014)	0.2669	0.2935
KTH-SE (2014)	0.7953	0.4607
HOUS-US (2014)	0.5323	-
second-best 2015	0.83	0.46
u-net (2015)	0.9203	0.7756

自分で実行した結果(1/2)

➤ モデルの正確さ(accuracy)

トレーニング画像：30枚細胞画像

エポック：300

認識率：0.8436

所要時間：29分54秒

```
292/300 [=====>.] - ETA: 43s - loss: 0.3562 - acc: 0.8423
293/300 [=====>.] - ETA: 38s - loss: 0.3559 - acc: 0.8425
294/300 [=====>.] - ETA: 32s - loss: 0.3556 - acc: 0.8426
295/300 [=====>.] - ETA: 27s - loss: 0.3553 - acc: 0.8428
296/300 [=====>.] - ETA: 21s - loss: 0.3550 - acc: 0.8430
297/300 [=====>.] - ETA: 16s - loss: 0.3548 - acc: 0.8432
298/300 [=====>.] - ETA: 10s - loss: 0.3545 - acc: 0.8434
299/300 [=====>.] - ETA: 5s - loss: 0.3544 - acc: 0.8435
300/300 [=====] - 1629s 5s/step - loss: 0.3543 - acc: 0.8436
```

➤ 実行結果

時間はかからないが、
認識率は高くない

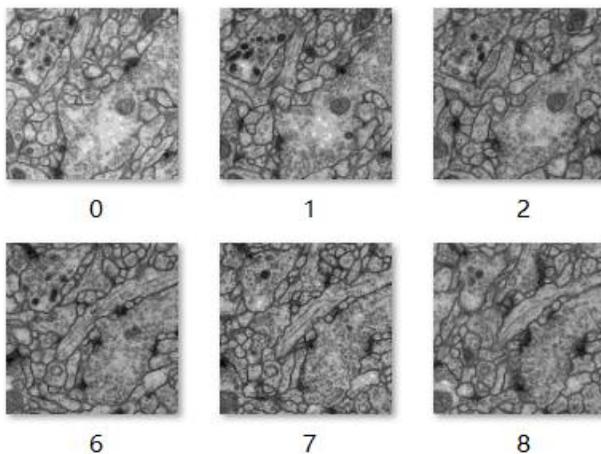


図1 テスト画像

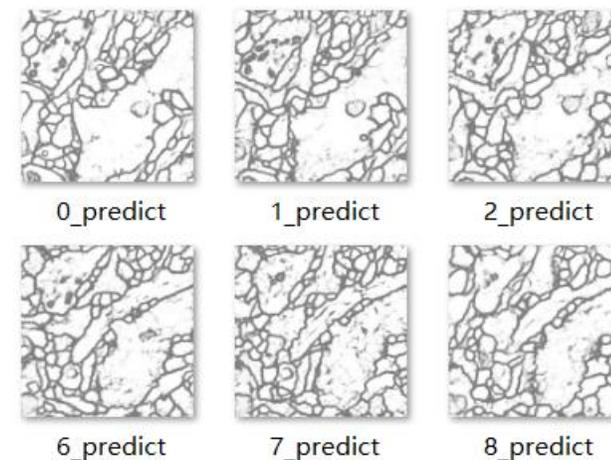


図2 実行結果

自分で実行した結果 (2/2)

➤ 他の画像実行結果

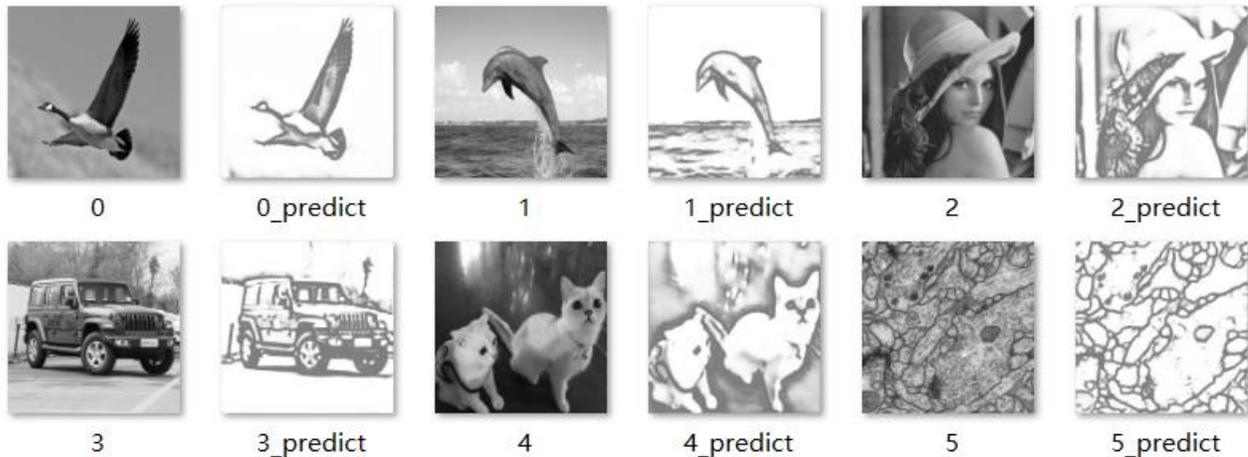


図1 物の画像の実行結果

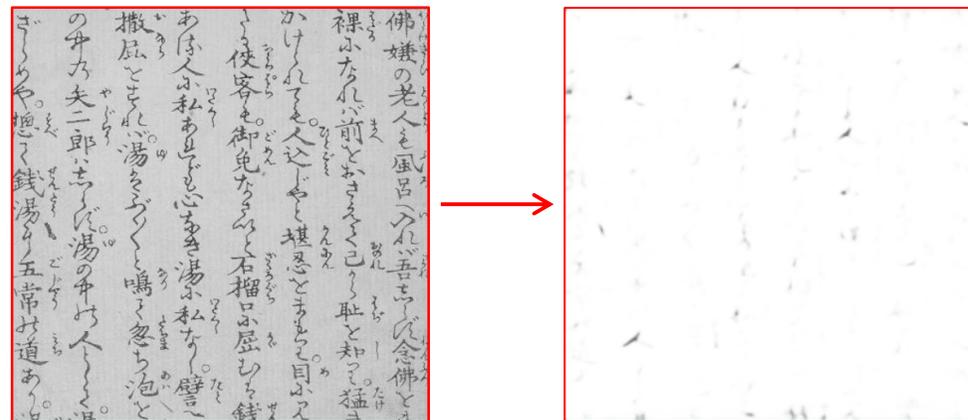


図2 文字の実行結果

物の輪郭を抽出は可能であるが、文字の抽出は困難

まとめ

U-Netは医療画像分割応用において非常に良い性能を得た

- 弾性変形のため、非常に少ないトレーニング画像が必要
- 出力されるのは単一のラベルではなく、2枚のフィーチャマップである
- 出力される画像は原図のサイズである
- トレーニング時間が少ない