

ビッグデータ活用に関する 新展開と国際連携 -医用画像処理の立場から-

森 健策

名古屋大学大学院情報学研究科 教授

名古屋大学情報基盤センター センター長

国立情報学研究所 医療ビッグデータ研究センター センター長



COVID-19 Cases – WHO report

Globally, as of 2:08pm CEST, 11 April 2021, there have been 134,957,021 confirmed cases of COVID-19, including 2,918,752 deaths, reported to WHO. As of 7 April 2021, a total of 689,248,795 vaccine doses have been administered.

Global Situation

134,957,021
confirmed cases

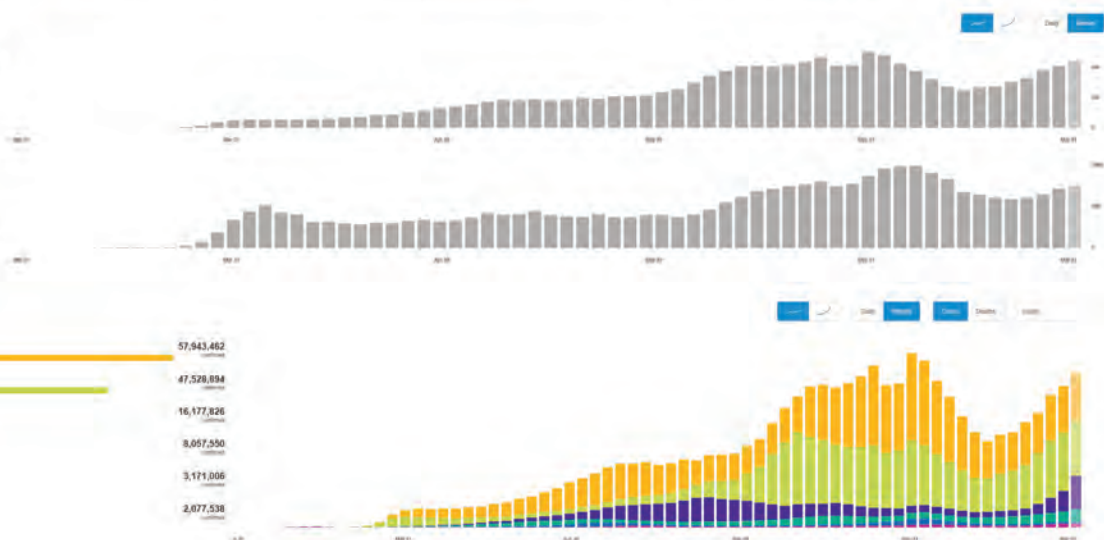
2,918,752
deaths

Source: WHO COVID Dashboard
© 2020 Nagoya University. All rights reserved.

Situation by WHO Region

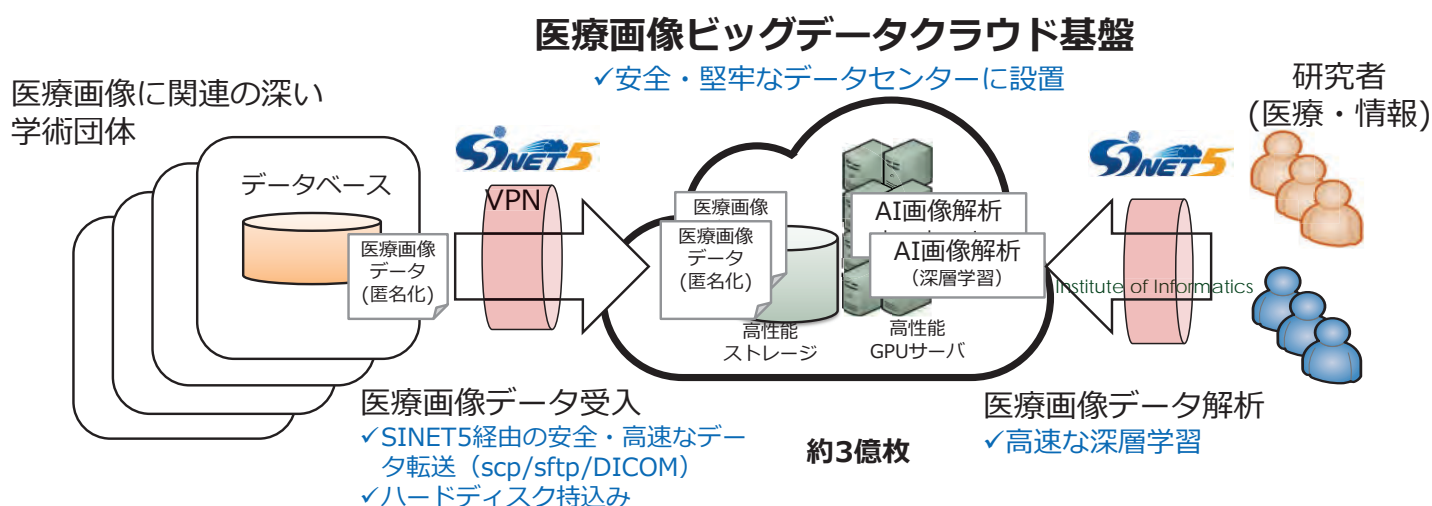
Americas	57,943,462
Europe	47,528,894
South-East Asia	16,177,826
Eastern Mediterranean	8,057,550
Africa	3,171,006
Western Pacific	2,077,538

Source: WHO COVID Dashboard
© 2020 Nagoya University. All rights reserved.



[WHO Coronavirus \(COVID-19\) Dashboard](https://covid19.who.int/)

医療ビッグデータ利活用を促進する クラウド基盤・AI画像解析に関する研究



COVID-19 パンデミック

- **COVID-19**
 - 感染者数 2.4億人 死者数 492万人 (2021/10/19) [1].
 - 予防法、診断法、治療法等の開発が重要
- **COVID-19 diagnosis**
 - RT-PCRが一般的に利用 (特異度は高いが感度が相対的に低い [2]) (42% to 71%).
 - CT画像による診断は比較的高い感度[3].
 - CT画像はCOVID-19を含むウイルス性肺炎の診断に有効

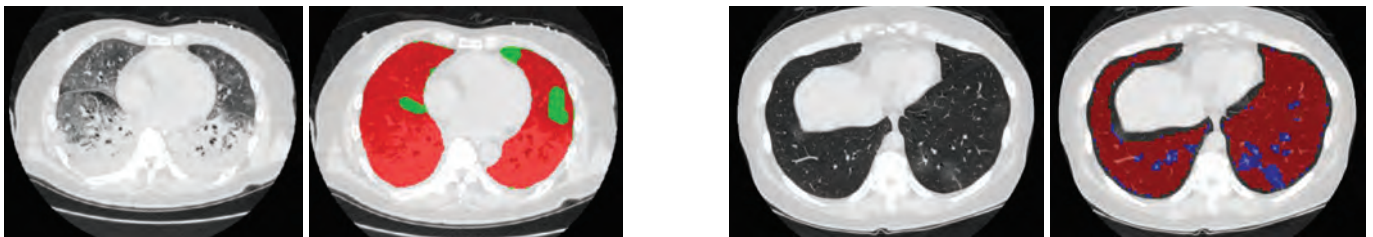
[1] "Coronavirus Update," <https://www.worldometers.info/coronavirus/>.

[2] Simpson S, et al., "Radiological Society of North America expert consensus document on reporting chest CT findings related to COVID-19: endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA," Radiology: Cardiothoracic Imaging 2(2) (2020).

[3] Ai T, et al., "Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," Radiology 296(2) (2020).

COVID-19のためのコンピュータ支援画像診断支援システム (CAD)

- 誤診断の防止と病院における読影負担軽減
 - 肺の定量的解析
- 要求される機能
 - 画像セグメンテーション、検出、分類、予測



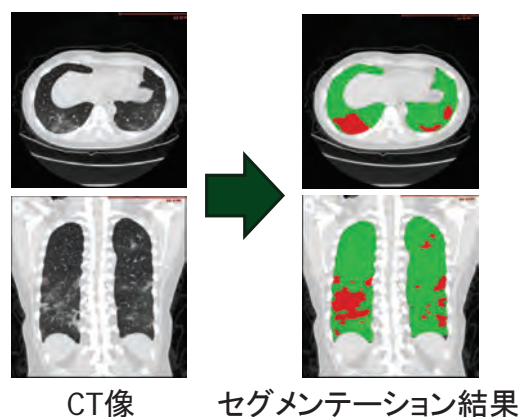
AIに基づいたCOVID-19 CT画像解析の例

- 肺野領域検出
 - 非常にシビアな肺炎の画像でも肺野領域を正しく認識できる
- 肺野領域分割
 - 肺野領域を正常, GGO, consolidated領域に分割
- 画像の分類
 - 入力CT画像をCOVID-19典型度に応じて分類する

COVID-19解析

肺野自動セグメンテーション

肺野領域の自動分割
画像所見の自動推定



7

肺野セグメンテーション

- 診断で用いるCT像の撮影条件は様々

撮影範囲の違い

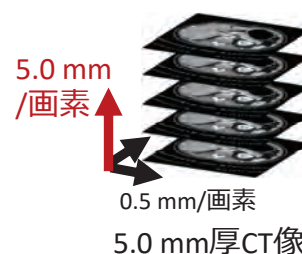
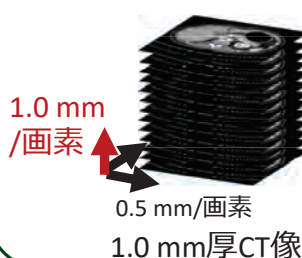


胸部のみのCT像



胸部と腹部を含むCT像

体軸方向の解像度の違い

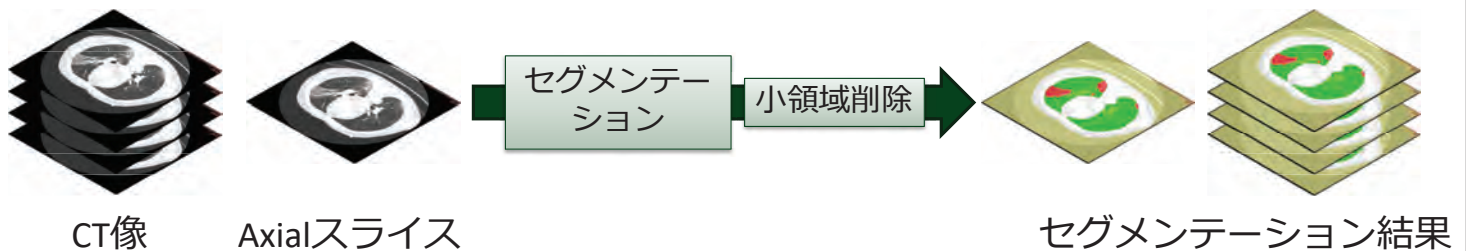


- AI画像診断支援の前処理としてCT像からの肺領域の識別は必須
- 様々な撮影条件への対応が必要

8

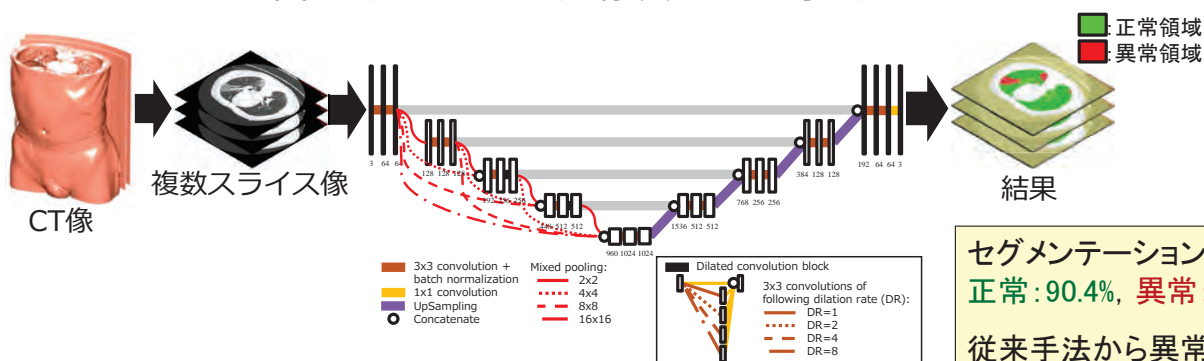
肺野領域自動セグメンテーション

- FCNを用いた自動セグメンテーション
 - 深層学習ネットワーク：2D U-Net（Axialスライス処理）
 - 損失関数にGeneralized Dice Lossを使用し小領域の見落とし防止
 - 学習データ：アノテーション付与したCOVID-19症例19例
 - 後処理として体積による小領域削除を実施



肺野領域自動セグメンテーション

- CT像から肺野の**正常**，**異常領域**を自動セグメンテーション
 - セグメンテーション用深層学習ネットワークを開発
 - 多様な形状と濃淡の異常領域に対応するためdense pooling, dilated conv.導入
 - スライス単位の処理により解像度の違い等に対応



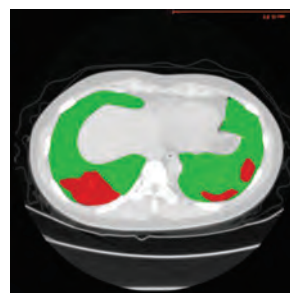
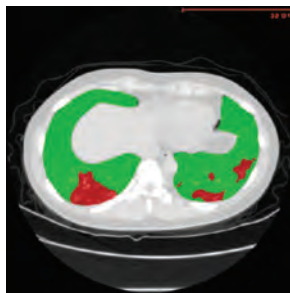
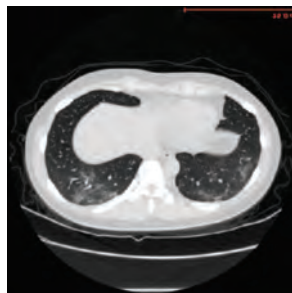
COVID-19用セグメンテーションNet

セグメンテーション精度 (DICE)

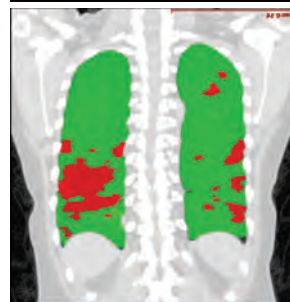
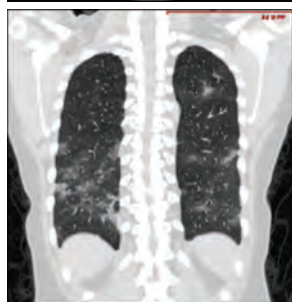
正常: 90.4%, 異常: 74.4%

従来手法から異常領域の精度向上
 (72.6% (Q Yan et al. 2020), 73.9% (D-P Fan et al. 2020))

自動セグメンテーション結果1



肺野正常
肺野炎症

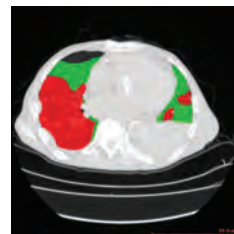
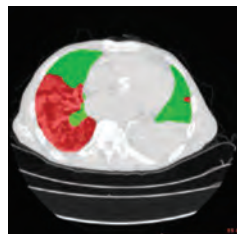
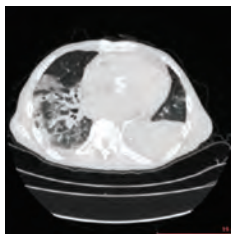


CT像

正解領域

自動セグメンテーション結果

自動セグメンテーション結果2



肺野正常
肺野炎症



CT像

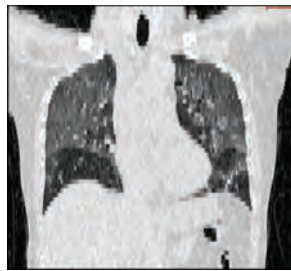
正解領域

自動セグメンテーション結果

自動セグメンテーション結果3 (5mm厚CT)



肺野正常
肺野炎症



CT像

正解領域

自動セグメンテーション結果

典型度・PCR検査結果別の肺野体積の分析

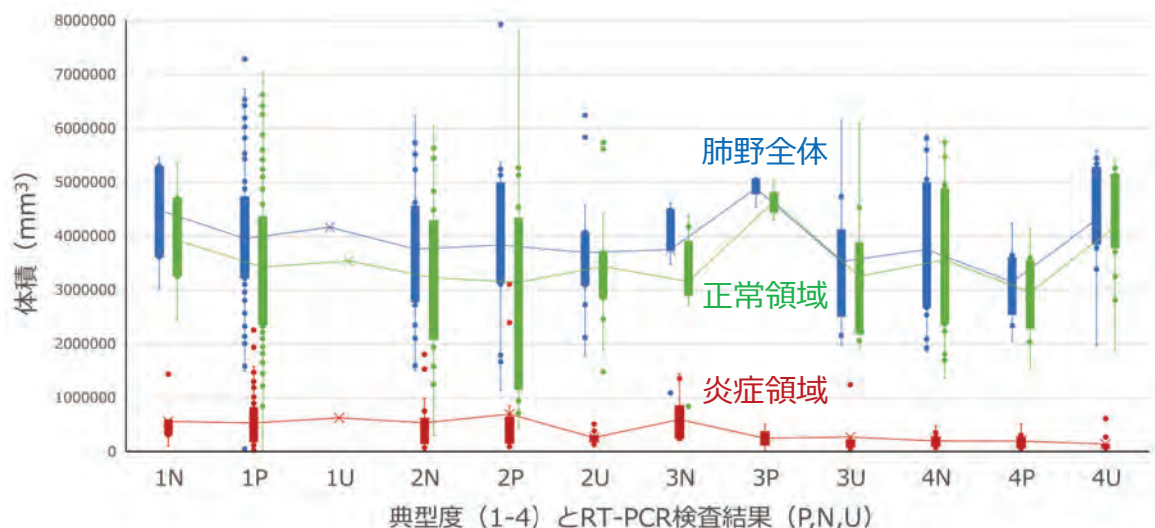
医師の所見と検査結果

COVID-19典型度

1. Typical appearance
2. Indeterminate appearance
3. Atypical appearance
4. Negative for pneumonia

RT-PCR検査結果

P : 陽性
U : 不明
N : 陰性



肺野体積の分析（炎症領域）

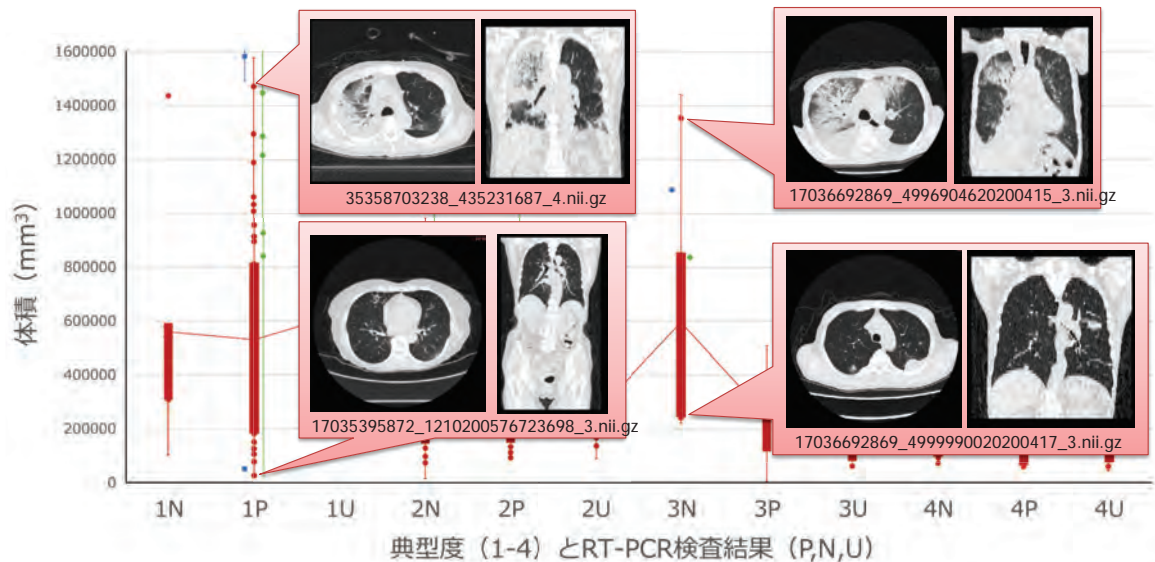
医師の所見と検査結果

COVID-19典型度

1. Typical appearance
2. Indeterminate appearance
3. Atypical appearance
4. Negative for pneumonia

RT-PCR検査結果

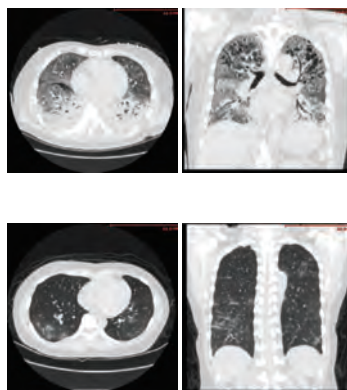
P : 陽性
U : 不明
N : 陰性



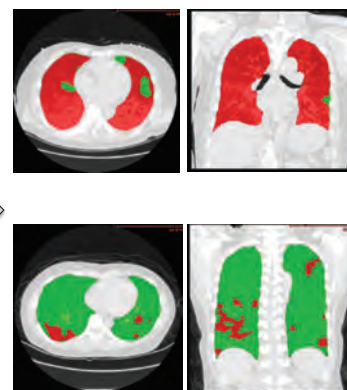
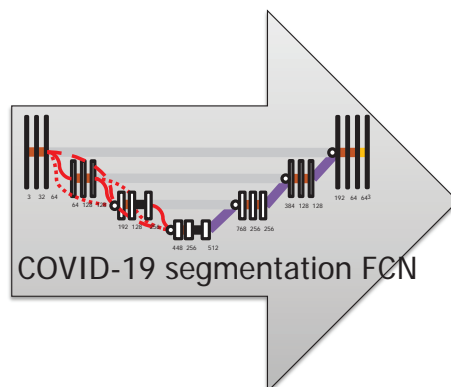
15

3次元画像としての処理

- CT像から肺野の**正常**，**異常陰影領域**を自動セグメンテーション
– 3Dパッチ画像ベースのセグメンテーション手法を作成



CT像

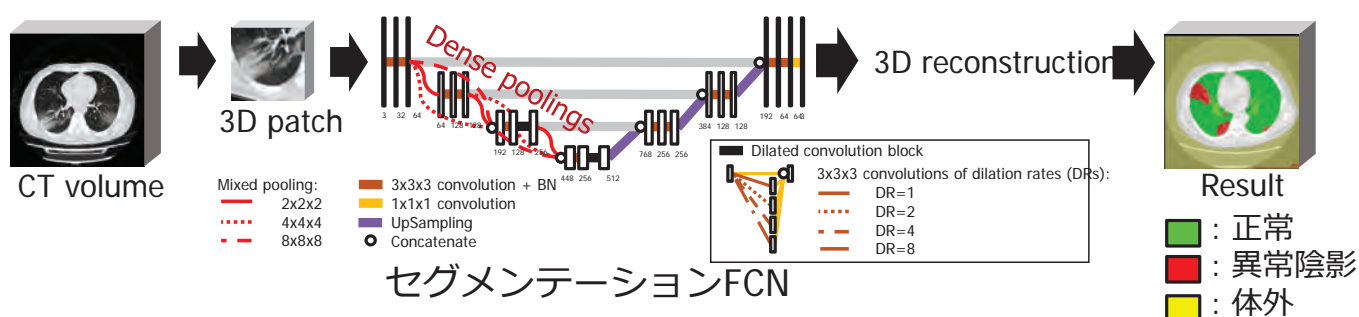


セグメンテーション結果

■ : 正常領域
■ : 異常陰影領域

3次元画像処理ネットワーク

- Encoder-decoder型のFCN
- 多様な形状と濃淡の異常領域に対応するため
dense pooling^[3], dilated conv.^[2]導入



[2] Yu F, Koltun V, Multi-scale context aggregation by dilated convolutions. International Conference on Learning Representations (ICLR), 2016

[3] Playout C, et al., A multitask learning architecture for simultaneous segmentation of bright and red lesions in fundus images. MICCAI 11071:101-108, 2018

3D手法のセグメンテーション結果

- 使用データ
 - COVID-19患者CT像20例
- FCNの学習
 - Dice loss使用
- 5分割交差検定で評価

セグメンテーション精度比較

2D/3D	手法	精度 (Dice)	
		正常領域	異常陰影領域
3D	提案FCN	0.864	0.744
	3D U-Net	0.840	0.732
	Yan et al. ^[1]	0.987	0.726
2D	FCN	0.904	0.744
	Inf-Net ^[2]	-	0.739
	2D U-Net	0.897	0.708

- 3D手法と
2D手法
どちらも高い精度を達成

[1] CT像731例で学習

[2] CT像45例で学習

[1] Q Yan, et al. "COVID-19 Chest CT Image Segmentation - A Deep Convolutional Neural Network Solution," arXiv:2004.10987, 2020.

[2] D-P Fan, et al. "Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images," IEEE Trans Med Imag, 39(8), 2020.

COVID-19解析

肺野自動セグメンテーション

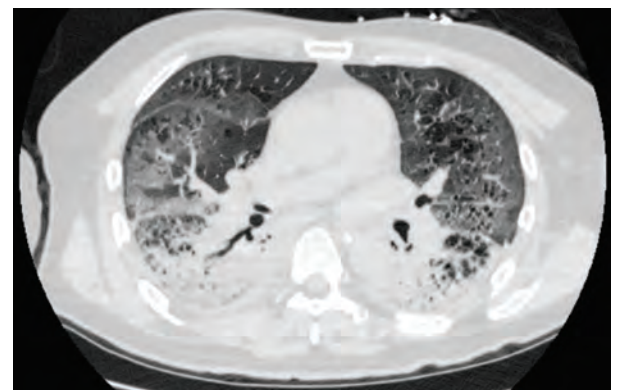
肺野領域の自動分割

画像所見の自動推定

19

肺野領域の分割

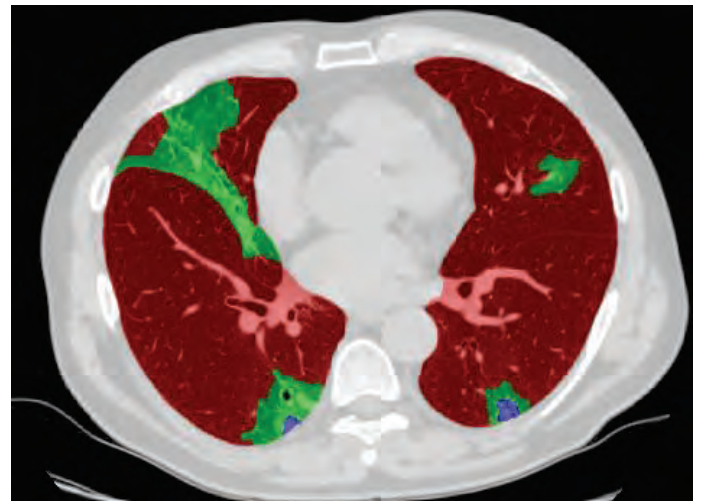
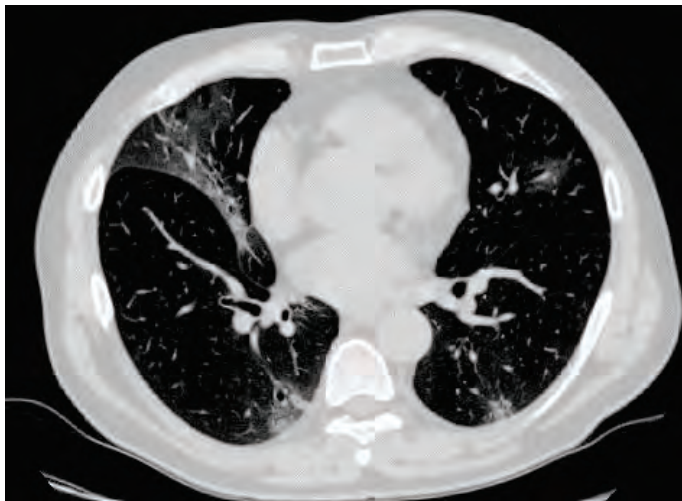
- 疾患状態の詳細解析には肺野内の陰影の位置・形状・サイズ等の情報が有用
 - 肺野内の状態の例
 - 正常部位
 - すりガラス状陰影
(ground glass attenuation)
 - コンソリデーション
- 肺野領域を自動分割
 - 教師なし学習を利用



COVID-19症例

20

肺野領域分割



■ Normal tissue
 ■ Ground glass opacity (GGO)
 ■ Consolidation

21

肺野領域の分割方法

- パッチベース画像処理とクラスタリングを用いた分割[1]
 - CT像の肺野領域をパッチ画像へと分割
 - Spherical K-meansを用いてパッチを特徴量へ変換
 - K-meansによるクラスタリング (3クラス)
 - クラスタリング結果に基づくCT像の領域分割



CT像

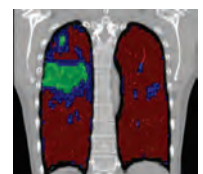
Spherical K-means

パッチ画像

パッチの
特徴量表現



K-Means
クラスタリング

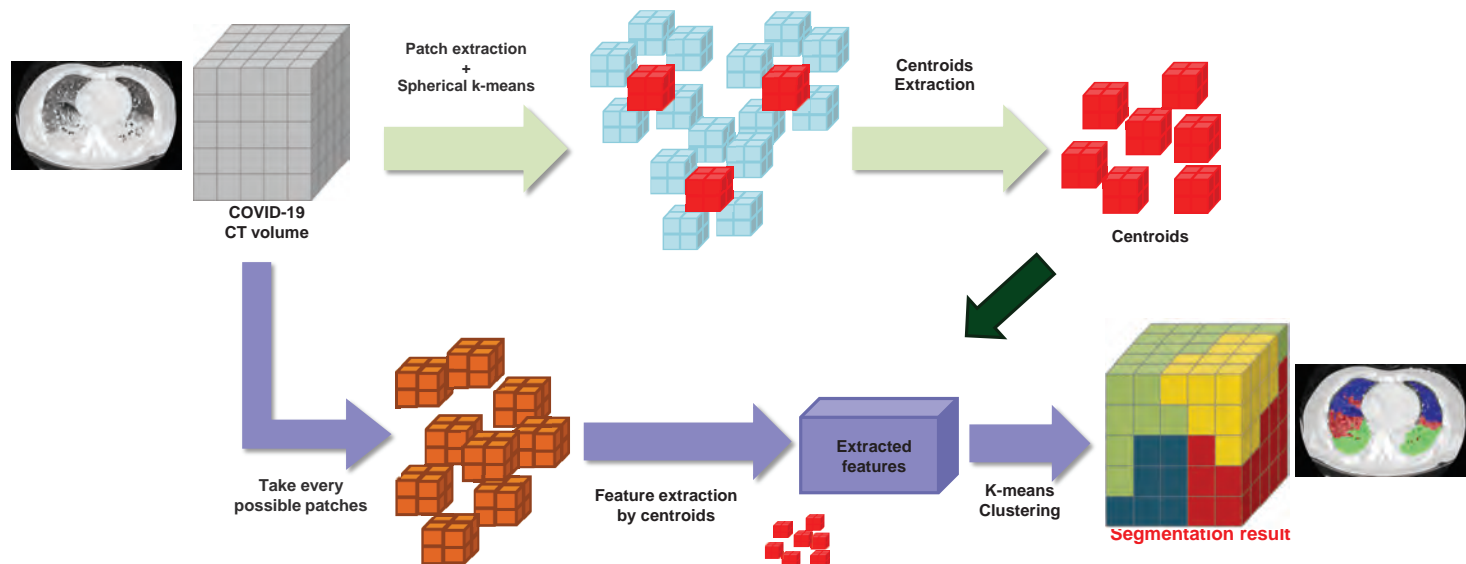


領域分割結果

[1] C Wang, T Moriya, Y Hayashi, M Oda, H Ohkubo, K Mori, Semi-supervised spherical K-means for segmenting idiopathic interstitial pneumonia from chest CT images, Int J CARS, 13(1), s27-28, 2018

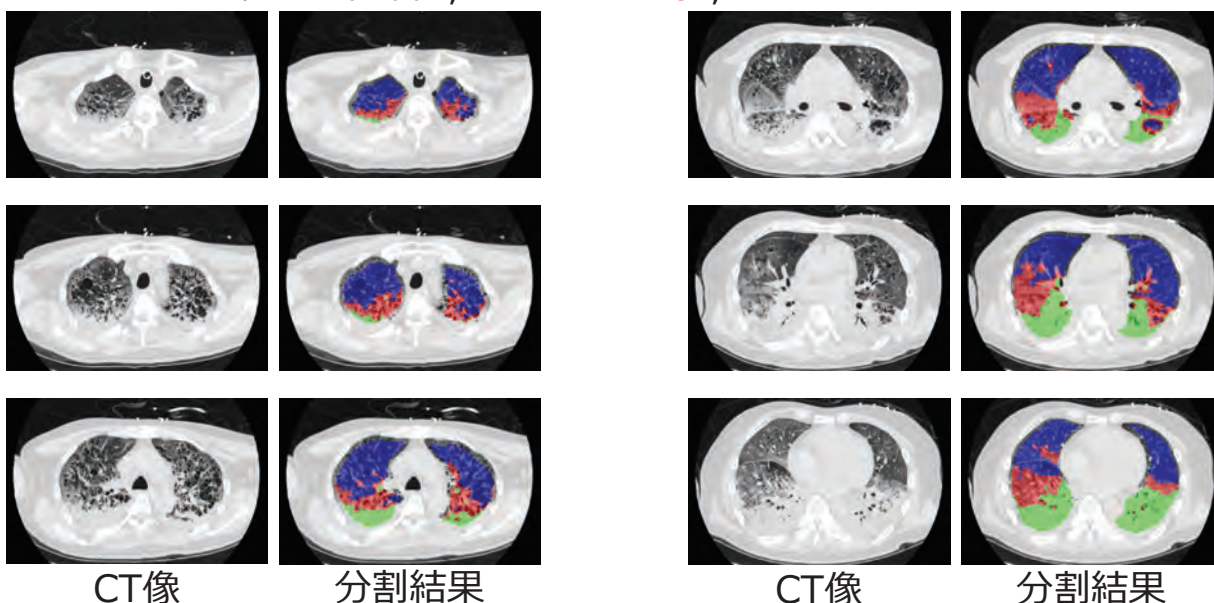
22

教師なし学習を利用した肺野領域を自動分割



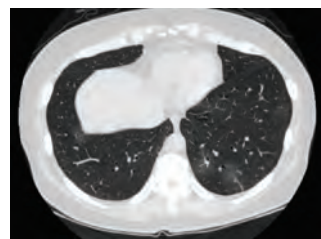
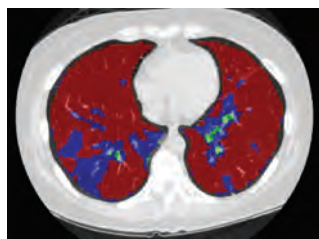
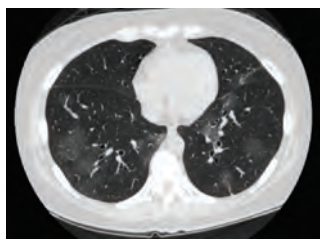
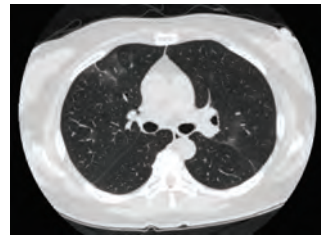
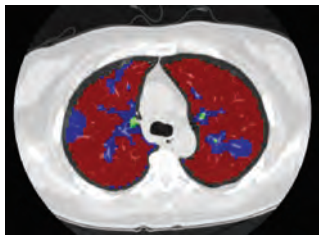
肺野領域分割結果1

色の意味 : 正常部位, すりガラス影, コンソリデーション



肺野領域分割結果2

色の意味：正常部位，すりガラス影，コンソリデーション



CT像

分割結果

CT像

分割結果

教師なし学習で肺野を自動分割可能

COVID-19解析

肺野自動セグメンテーション

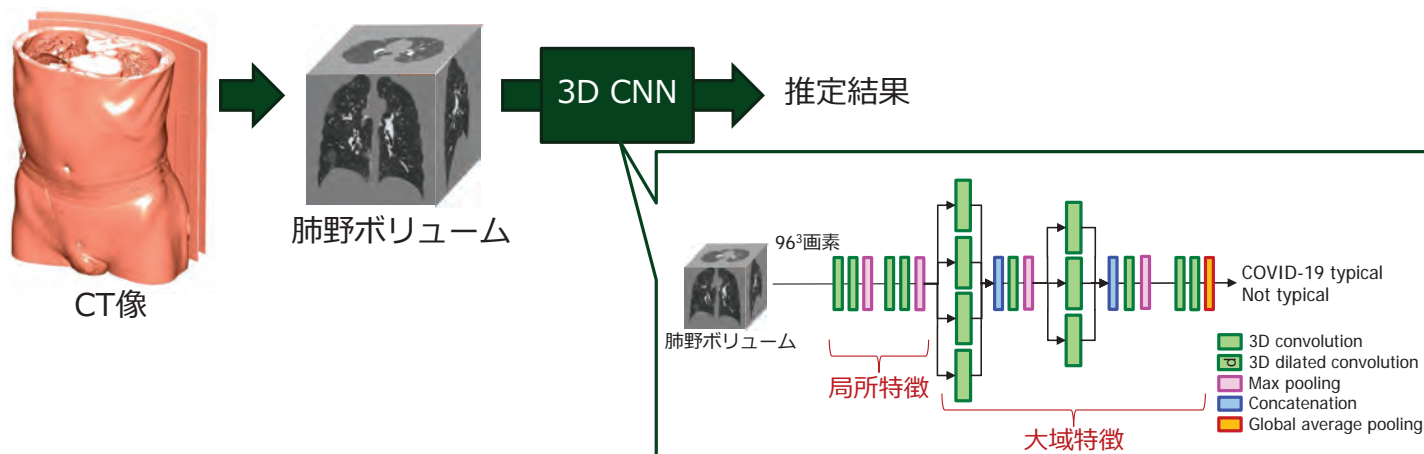
肺野領域の自動分割

画像所見の自動推定

3D CNNを用いた画像所見推定

• 推定手法

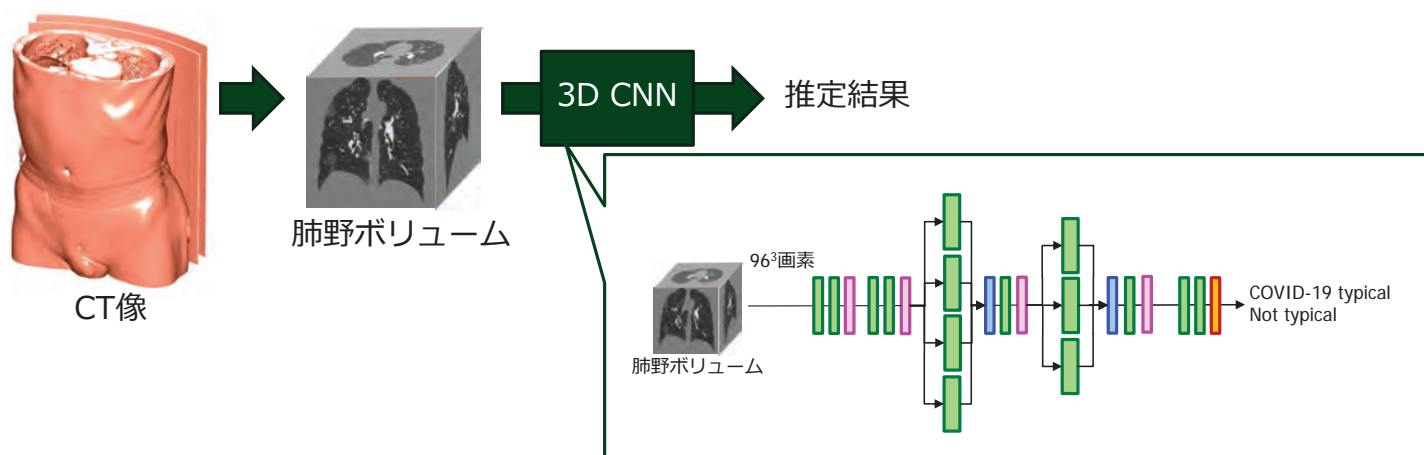
- CT像から取り出した肺野ボリュームから3D CNNで推定実施
 - 3D CNNにDilated convolutionを使用し肺野の局所及び大域特徴を判断に利用



3D CNNを用いた画像所見推定

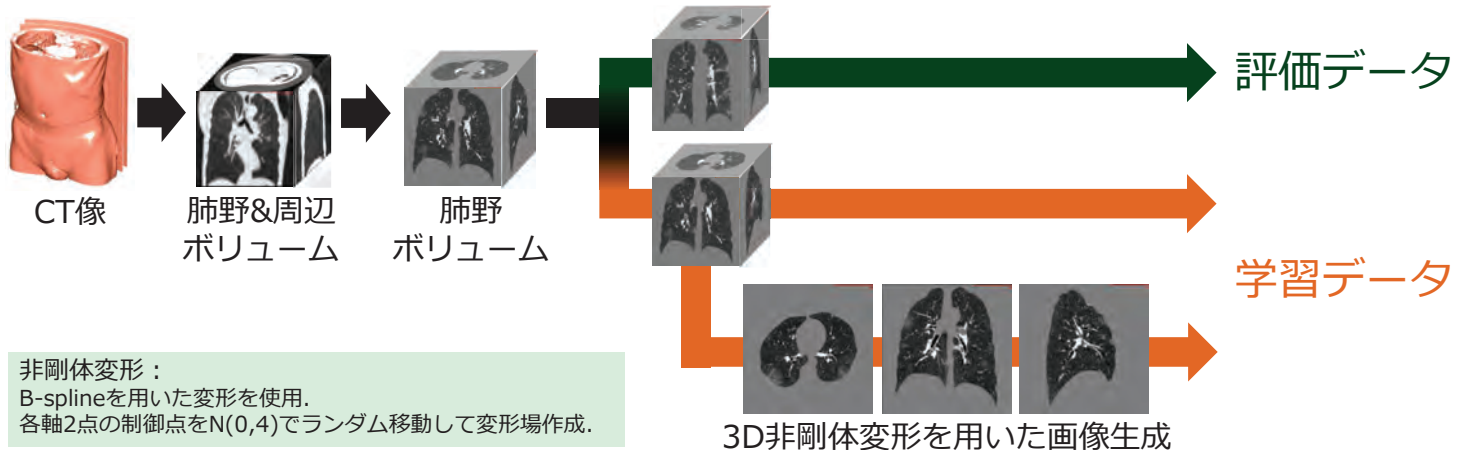
• 推定手法

- CT像から取り出した肺野ボリュームから3D CNNで推定実施
 - 3D CNNにDilated convolutionを使用し肺野の局所及び大域特徴を判断に利用



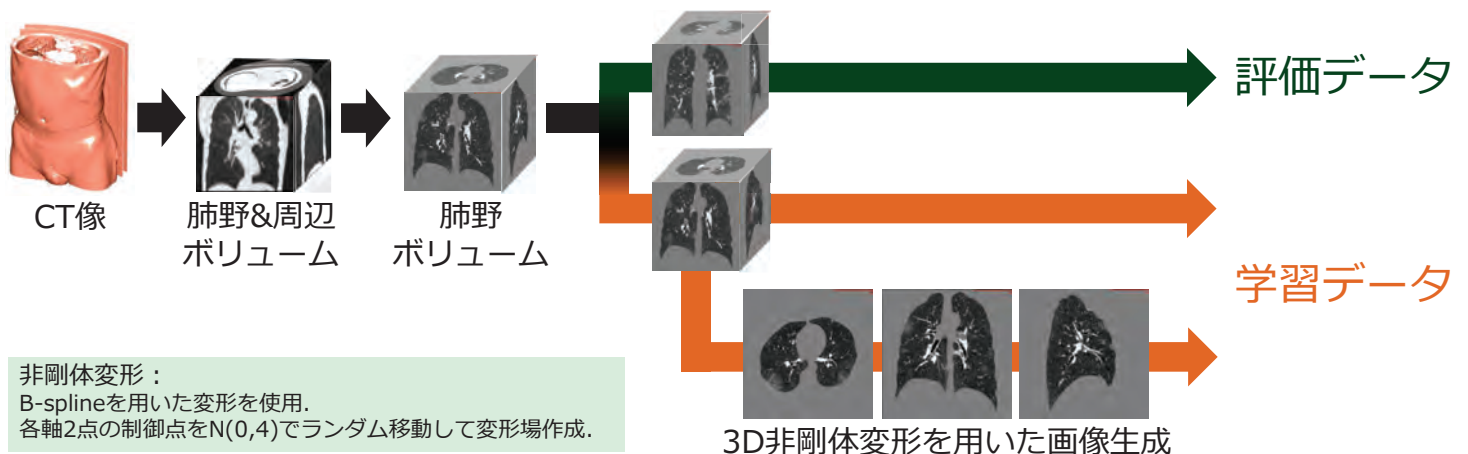
3D CNNを用いた画像所見推定

- 学習用肺野ボリューム生成
 - より多くの肺野形状バリエーション学習のため
3D非剛体変形処理を用いて肺野ボリューム変形



3D CNNを用いた画像所見推定

- 学習用肺野ボリューム生成
 - より多くの肺野形状バリエーション学習のため
3D非剛体変形処理を用いて肺野ボリューム変形



画像所見推定の結果

- COVID-19典型度「高（1,2）」vs「低（3,4）」の2分類

- データ

- 評価に「1,2」15例・「3,4」15例
 - 学習に787例（元画像）+ 1600例（生成画像）

- 分類精度83.3%

COVID-19典型度

1. Typical appearance
2. Indeterminate appearance
3. Atypical appearance
4. Negative for pneumonia

		推定結果	
		高（1,2）	低（3,4）
正解	典型度高（1,2）	14	1
	典型度低（3,4）	4	11

分類精度

93.3%

73.3%

- COVID-19典型度の高い症例を高精度に発見可能

COVID-19禍でのオープンデータ・オープンサイエンス

- COVID-19禍において大量のデータがオープンにアクセスできるようになった
 - 学术界が迅速に対応
- オープン化の対象
 - 論文
 - データ (遺伝子配列、画像データ)
- 医用画像の場合
 - RSNA COVID-19画像データベース

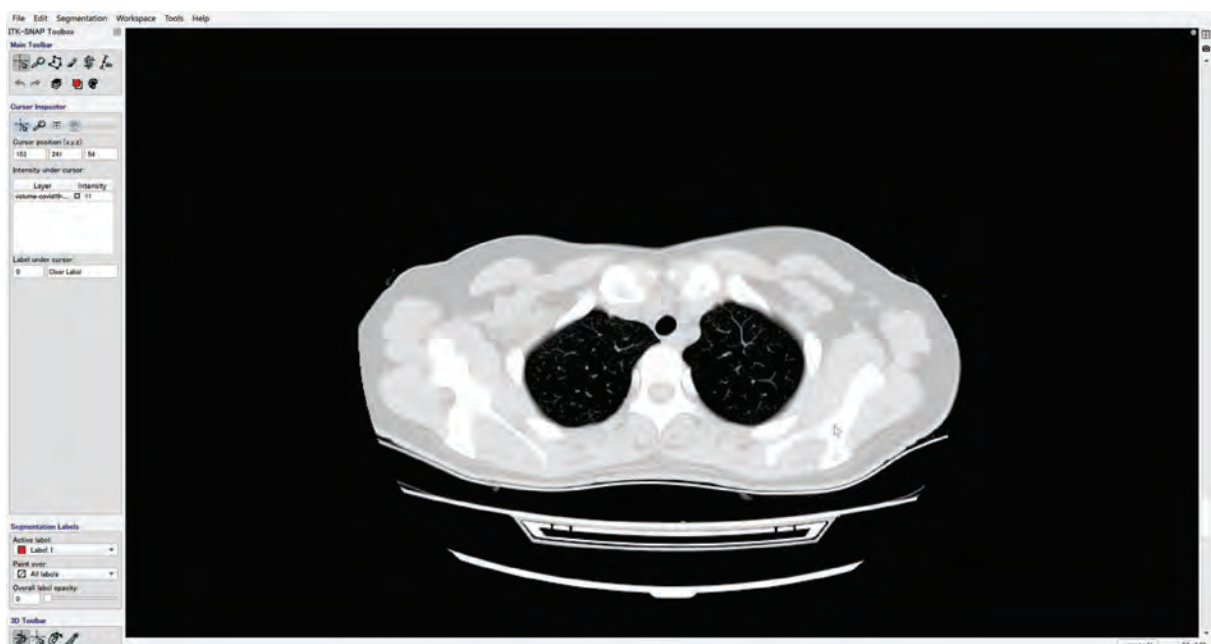
MICCAI COVID-19-20 Challengeのホームページから

- COVID-19 Lung CT Lesion Segmentation Challenge - 2020
 - The COVID-19-20 challenge will create the platform to evaluate emerging methods for the segmentation and quantification of lung lesions caused by SARS-CoV-2 infection from CT images. The images are multi-institutional, multi-national and originate from patients of different ages, gender and with variable disease severity.
- <https://covid-segmentation.grand-challenge.org/Final-Ranking/>



35

MICCAI COVID-19-20 Challenge画像



36

MICCAI COVID-19-20 Challenge

- COVID 19-20 CT画像
 - 学習 199 cases
 - 評価 50 cases
 - テスト 46 cases
- 私の研究室からも参加
- Team “Zheng + Zhang + Oda”
 - Ensemble 2D nnU-Net + 3D low-res nnU-Net
 - 世界ランク9位

Grand Challenge: Authors Reader Studies Challenges Algorithms Products

Challenges / COVID-19 Lung CT Lesion Segmentation Challenge - 2020 / Challenge Final Ranking

Congratulations to all the participants in the challenge!

Top 10 consensus ranking across tasks

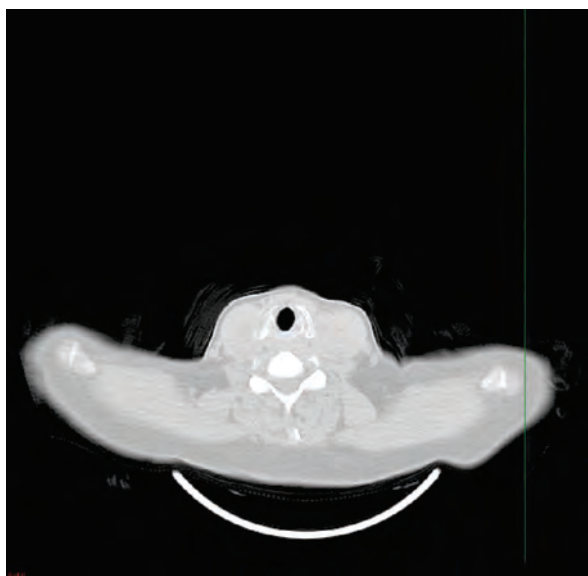
Rank	Algorithm Name	Rank User	Authors	Affiliation
1	3.383	1st place	Wenjun Hu, Jiaojiao Zhang and Hong Xia	Northeastern Polytechnical University, China
2	6	2nd place	Fabrizio Netti, Peter M. Fu, Michael Glick, Tobias Hengstler, Klaus H. Maier	Division of Medical Image Computing, German Cancer Research Center, Germany
3	7.687	3rd place	Chen Tang	Lundbeck High School, USA
4	8.417	4th place	Qing Yu, Qikai Li, Kang Deng	SJTU, China
5	8.5	5th place	Jae Seon, Donghye, Pooja, Mohit, Boudadi, Mouna, Benou, Vignje, Boudi, Garcia, Santa Cruz, Andreu, Hous	Univ. of Luxembourg, Luxembourg; Centre for Systems Biomedicine, Luxembourg
6	9.167	6th place	Song Zheng, Luoma Zheng	Nagoya University, Japan
7	9.167	7th place	Vitali Lashin	United Institute of Informatics, Ploiesti (IIRPI), Romania



MICCAI COVID-19 Challenge セグメンテーション結果



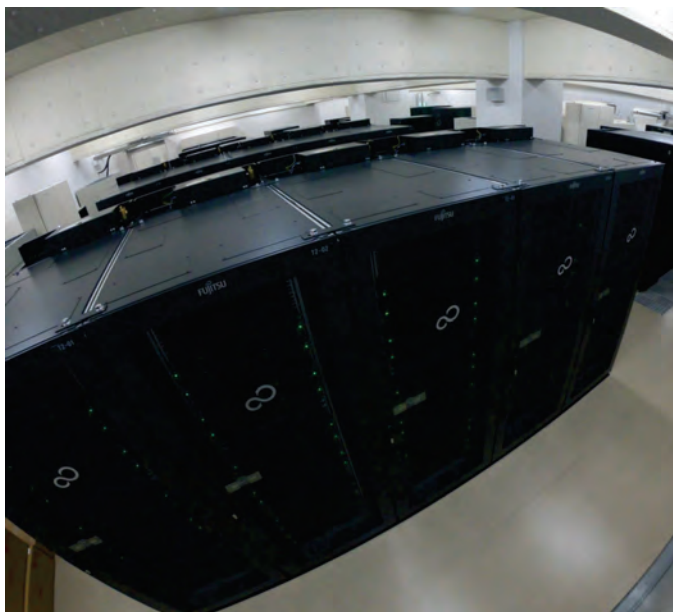
原画像



セグメンテーション結果

スーパーコンピュータ「不老」について

不老



40

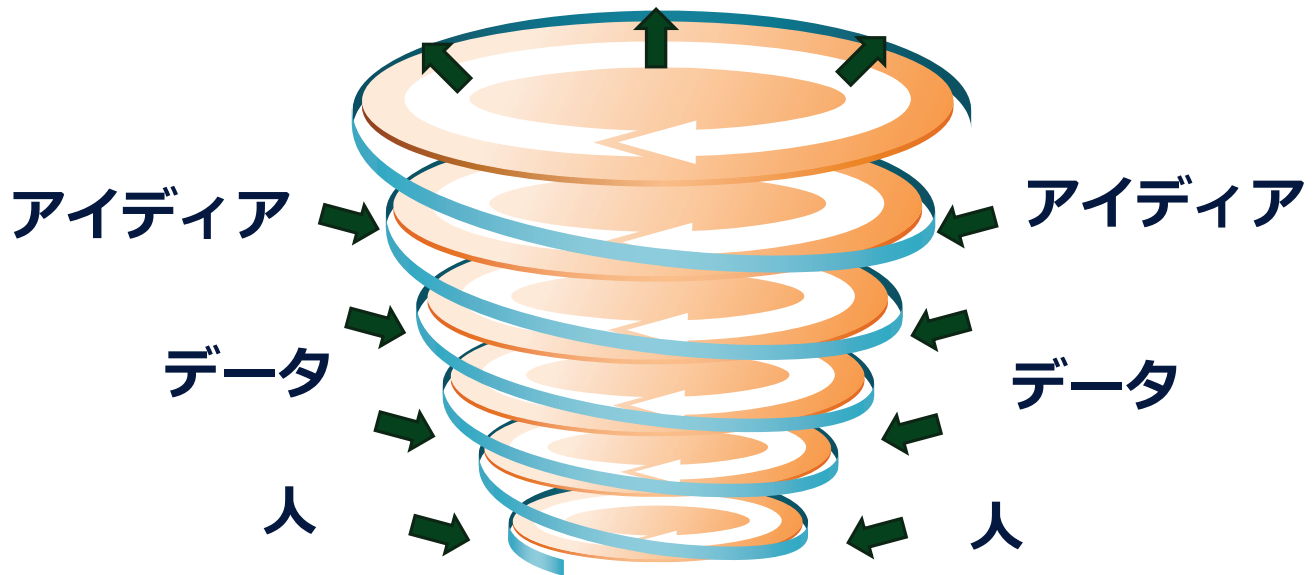
スーパーコンピュータの特徴

- 極めて高速なコンピュータ
- 非常に数多くのコンピュータ(CPU)を超高速ネットワークで接続することで高速演算を実現
- スーパーコンピュータが持つ性能を最大限に生かすにはソフトウェアでの工夫が必要



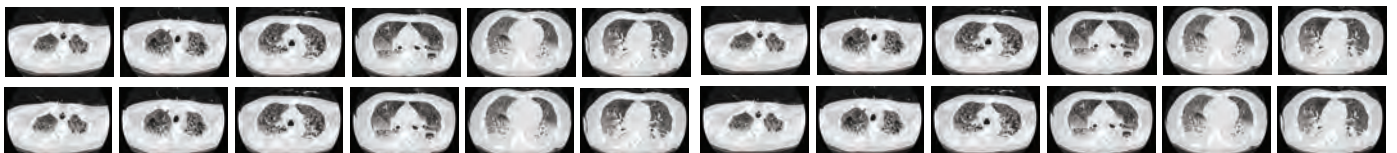
41

大量の計算能力はよい手法を実現するうえで必須! 優れた手法

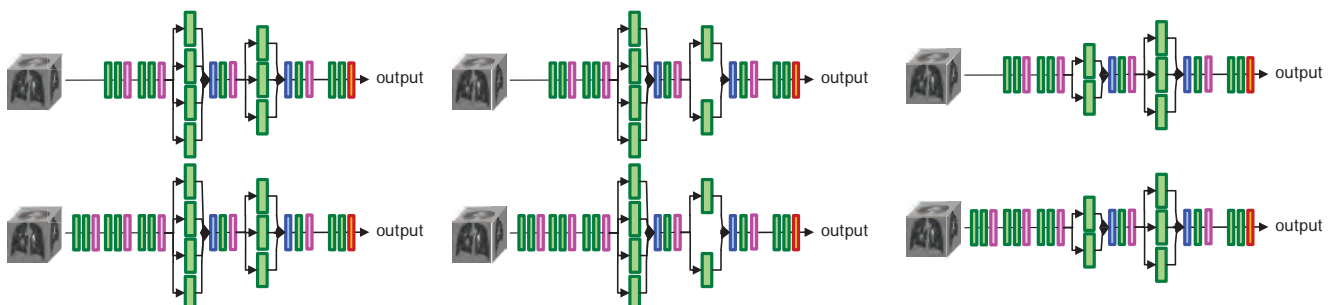


AI計算におけるスパコンの重要性

- 大量の画像を学習させる必要がある

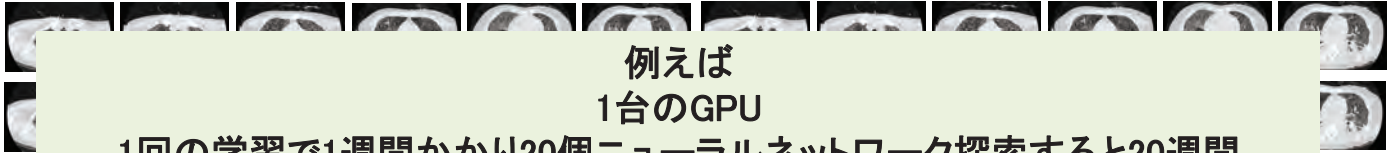


- 大量の計算で最適なニューラルネットワーク構成を求める



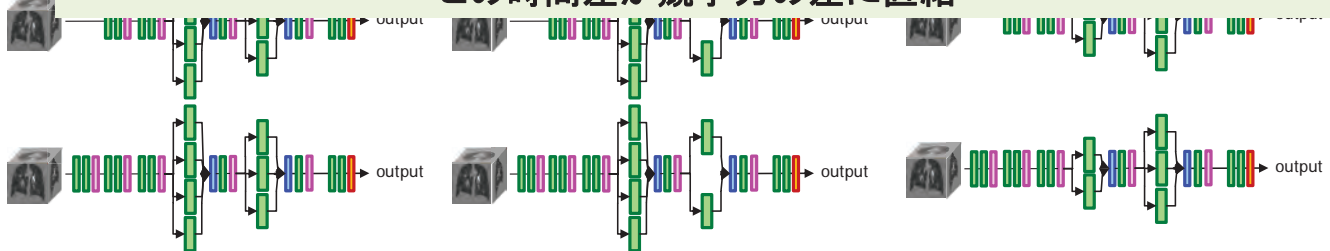
AI計算におけるHPCの重要性

- 大量の画像を学習させる必要がある



例えば
1台のGPU
1回の学習で1週間かかり20個ニューラルネットワーク探索すると20週間
1000台のGPUだと $20/1000 = 0.02$ 週間 = 0.14日

この時間差が競争力の差に直結



UK biobank

- 研究者が自由にアクセスできるデータベースの存在がAI研究を加速
- ナショナルDBを利用していろいろな知識発見や技術開発

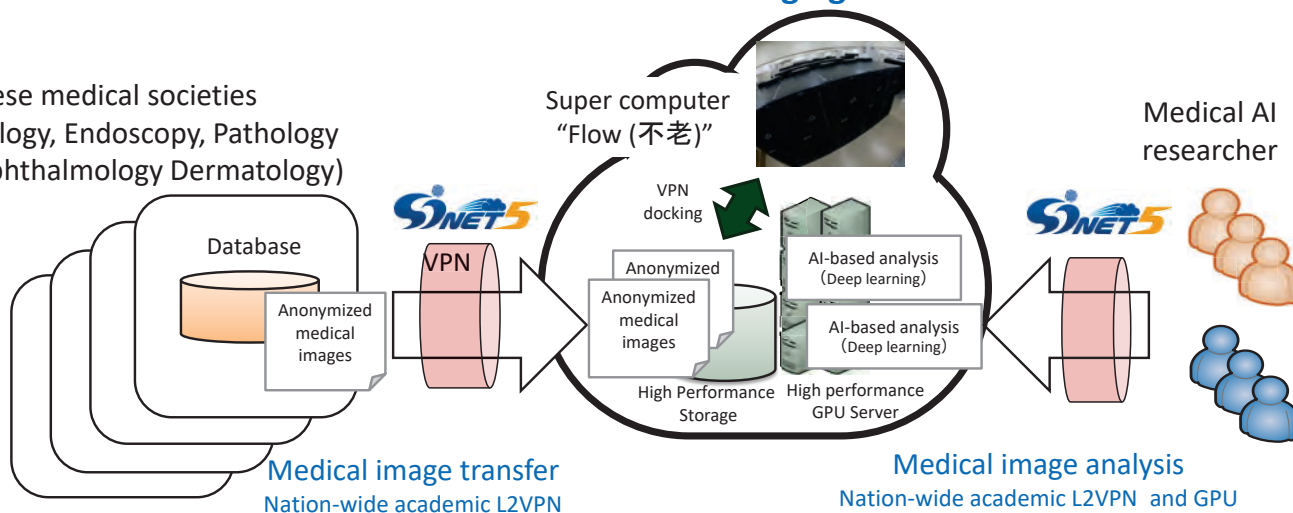


<https://www.ukbiobank.ac.uk/>

医用画像データベースプロジェクトと スパコンドッキング

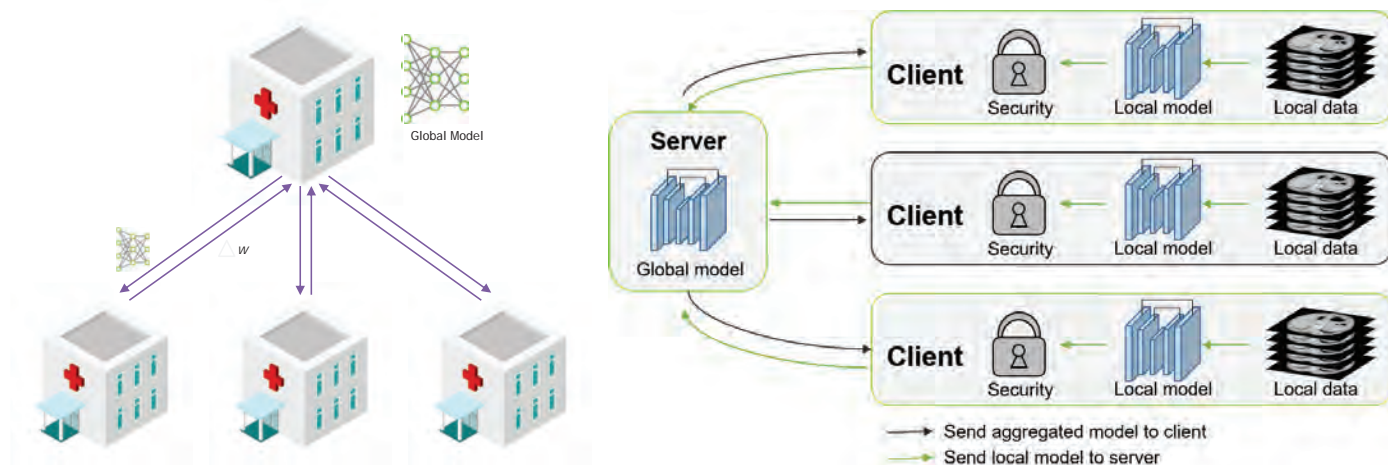
Medical Imaging Cloud

Japanese medical societies
(Radiology, Endoscopy, Pathology
US, Ophthalmology Dermatology)



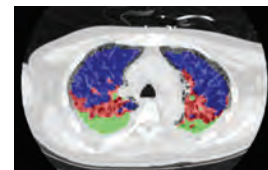
連合学習

- 機械学習モデルのみ共有・データはシェアしない



むすび

- **ビッグデータ活用に関する新展開と国際連携**
 - 医用画像処理の立場からデータを活用したCOVID-19 AIの開発
- **COVID-19 CT像AI**
 - 肺野領域セグメンテーション、肺野領域分割、典型度診断
- **クラウド基盤データ利活用**
- **国際連携**
 - オープンデータ
- **スパコンを用いたAI開発**



むすび