

2020年（令和2年）9月28日

## 新型コロナウイルス肺炎 CT 画像を AI 解析するためのプラットフォームを開発 全国の病院から集めた CT 画像を AI で選別し高品質な AI 研究用データセットとして整備

情報・システム研究機構 国立情報学研究所（<sup>エヌアイアイ</sup>NII、所長：喜連川 優、東京都千代田区）は、東海国立大学機構 名古屋大学（総長：松尾 清一、愛知県名古屋市）、順天堂大学（理事長：小川 秀興、東京都文京区）、日本医学放射線学会（理事長：青木 茂樹、東京都文京区）などと共同で、新型コロナウイルス肺炎（COVID-19 肺炎）の CT 画像を AI 解析するためのプラットフォームを開発しました。

今回の AI 解析用プラットフォームの開発では、NII の医療ビッグデータクラウド基盤（図 1）に収集・集積した 1 億 6 千万枚を超える CT 画像を活用し、大量の画像の中から肺炎 CT 画像を選別する機械学習手法をまず開発しました。そして、その新たな選別手法によりリスト化された肺炎 CT 画像に、実際の PCR 検査結果と放射線医による COVID-19 肺炎典型度<sup>(\*1)</sup>の判定結果を付与してデータベース化しました。このような手順で質の高い AI 向け学習データセットを整備し、COVID-19 肺炎研究のための新たな AI 解析用プラットフォームとして開発しました（図 2）。

この COVID-19 肺炎症例 CT 画像の AI 解析プラットフォームはすでに COVID-19 肺炎研究に活用されています。名古屋大学の研究チームは COVID-19 肺炎症例 CT 画像データベースに独自に AI を適用することで、精度の高い COVID-19 肺炎典型度判定の手法を開発しました（図 3）。この手法では、COVID-19 肺炎の典型例の識別タスクで 83.3%の典型度識別性能を達成しています<sup>(\*2)</sup>。また、炎症などの影響により CT 画像上で肺が識別困難な場合に、AI が的確に肺の形状を推定する手法も開発しています。

NII を中心とした研究チームでは、COVID-19 肺炎研究のための AI 解析用プラットフォームをさらに整備するとともに、他の研究チームと連携して AI アルゴリズムを改良し、AI 選別や判定の精度向上を目指します。

### 【背景】

2019 年末ごろに発生した新型コロナウイルス感染症（COVID-19）は急速に世界に広まり、多数の患者が発生しています。2020 年 9 月下旬における世界の累計感染者数は約 3,200 万人、死亡者数は約 98 万人となっています<sup>(\*3)</sup>。我が国の累計感染者数は約 8 万人、死亡者数は約 1,500 人<sup>(\*4)</sup>で、さらに増加を続けています。COVID-19 患者に対し適切な治療を行い、他者への感染を防ぐ上で、正確な診断方法が必要とされています。COVID-19 の検査では RT-PCR（reverse

transcriptase polymerase chain reaction) が多く用いられますが、その Sensitivity (感度) は 42% から 71% と低い値を示します(\*5)。これに対し CT 画像による COVID-19 検査の Sensitivity (感度) は 97% と高いといった特徴があります(\*6)。日本は人口あたりの CT 撮影装置保有数が高いため、COVID-19 検査において CT 画像を活用できると考えられます。CT 画像を用いた医師の診断を支援し、判断の定量化を目指す上で、コンピューターにより肺の状態を客観的に判断する評価手法の確立が求められています。

### 【開発した技術：AI 解析用プラットフォーム】

NII は、東海国立大学機構 名古屋大学、順天堂大学、日本医学放射線学会などと共同で、新型コロナウイルス肺炎 (COVID-19 肺炎) CT 画像を迅速に AI 解析するプラットフォームを開発・整備しました。このプラットフォーム開発は、主に次の 2 つの技術開発からなっています。

- (1) COVID-19 肺炎 CT 画像を機械学習で効率良く選別する手法
- (2) 放射線医による判定と PCR 検査結果が付帯した高品質学習データの整備

このプラットフォーム開発には、NII が中心となって構築・運営している医療画像ビッグデータクラウド基盤を活用しています。このクラウド基盤には、匿名化した CT 画像を全国の病院から日本医学放射線学会を通じて収集し、過去 2 年間に 1 億 6 千万枚を超える CT 画像を蓄積しています (図 1)。

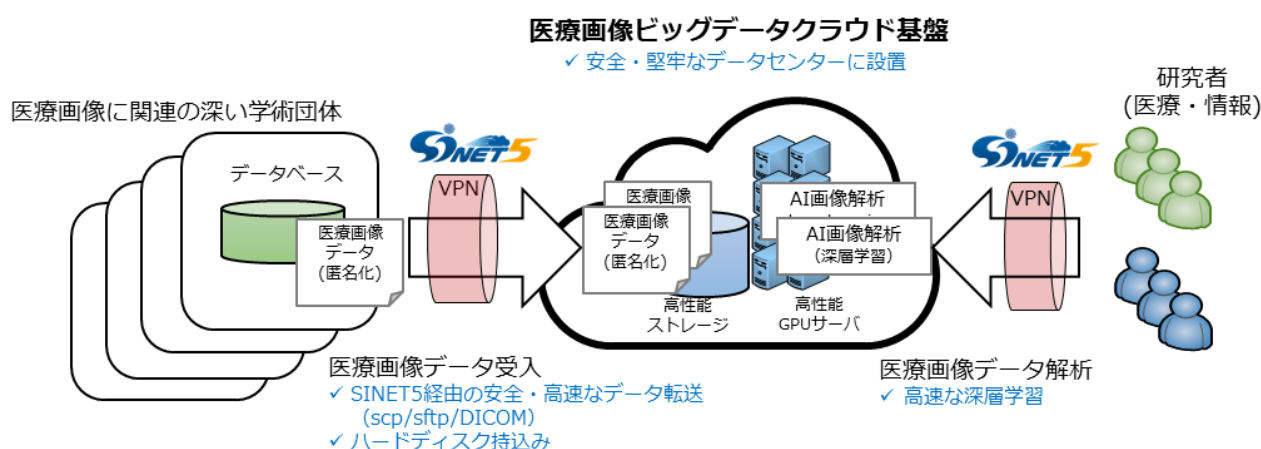


図 1 医療ビッグデータ利活用を促進するクラウド基盤のしくみ

まず、これらの CT 画像から COVID-19 パンデミック以前と以後の肺炎症例を機械学習によって選別する手法を開発しました (図 2)。この手法を適用した結果、クラウド基盤から非 COVID-19 肺炎 367 症例、COVID-19 肺炎 320 症例 (2020 年 9 月 8 日現在) が抽出されました。この抽出された CT 画像に PCR 検査結果と放射線医が判定した COVID-19 肺炎典型度のメタデータを付与しました。このメタデータの付与では特に順天堂大学、日本医学放射線学会と密接に連携し、AI 機械学習データとして質の高い COVID-19 肺炎 CT 画像をデータベースとして整備し、AI 解析用プラットフォームとして開発しました。

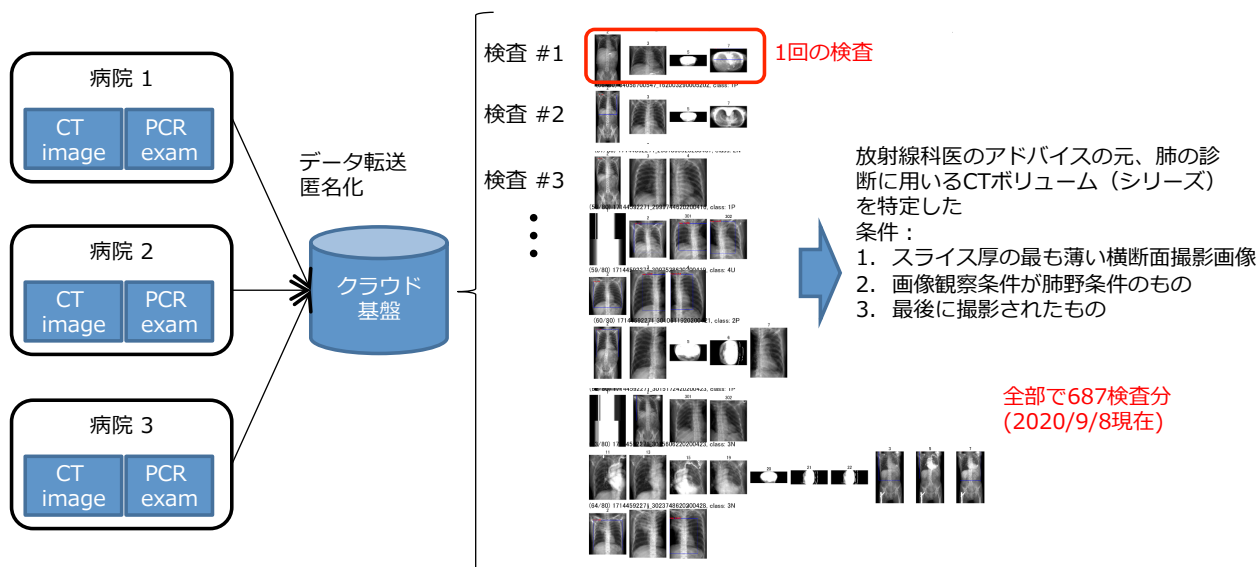


図2 COVID-19 肺炎 CT 画像データベースの構築

### 【AI 解析用プラットフォームの活用事例】

NII を中心としたチームが開発した AI 解析用プラットフォームは、すでに COVID-19 関連の研究に活用されています。名古屋大学の研究チームは、この COVID-19 肺炎 CT 画像データベースに対して AI を適用し、高い精度で COVID-19 肺炎の典型度を判定できる AI を開発しました(図 3)。

色の意味：正常部位、すりガラス影、コンソリデーション

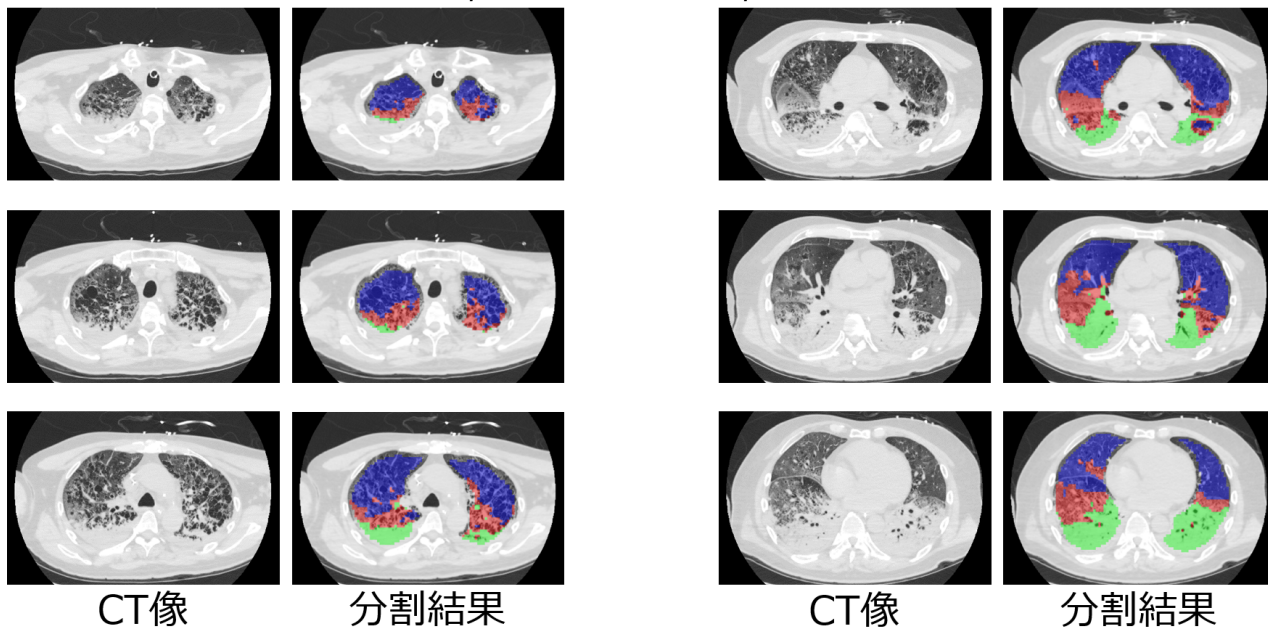


図3 COVID-19 肺炎典型度を判定する AI による部位の分割結果。空気に富む肺胞領域は CT では X 線透過率が高いために黒っぽく映る（正常部位）。ここに炎症が起こると X 線透過率が低下してすりガラスのような陰影を呈する（すりガラス影）。さらに細胞浸潤が激しくなると、いっそう X 線透過率が低下して白く一様な陰影となる（コンソリデーション）。

名古屋大学チームの AI は、胸部 CT 画像を入力すると、その CT 画像の COVID-19 肺炎典型度を自動判定することがきます。この名大チーム AI は、COVID-19 肺炎の典型例とそうでない例の識別タスクにおいて、現時点で 83.3%程度の典型度識別性能を達成しました。また、この解析を可能とするために、炎症などの影響で CT 画像上の肺の形状が識別困難な場合でも、AI が的確に肺の形状を推定できる手法も開発しました。なお、この名古屋大学チームを中心とした研究成果は、2020 年 9 月に開催された第 39 回日本医用画像工学会大会にて発表されました。

### 【社会的インパクトと今後の展望】

医療画像 AI の研究開発では、AI を学習させるために、一病院だけではなく、様々な病院で取得された多様で高品質な画像データ群が極めて重要です。そのため、今回の COVID-19 肺炎の CT 画像データベース構築と解析プラットフォームの開発では、NII と日本医学放射線学会をはじめとする 6 つの医療画像に関連する臨床学会が中心となって進めAMED が支援する研究プロジェクト「臨床研究等 ICT 基盤構築・人工知能実装研究事業」の「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」<sup>(\*)7)</sup>を活用しました。

この研究プロジェクトでは COVID-19 の顕在化以前から、匿名化等の処理を行った CT 画像を含む医療画像を、国内の病院等から NII が運営する医療ビッグデータクラウド基盤へ転送し蓄積していました。この転送には、NII が運営する超高速の学術情報ネットワーク SINET5<sup>(\*)8)</sup>を利用しており、さらに、画像処理研究者もこのクラウド基盤に SINET5 でアクセスして医療画像 AI の研究開発を行っています（図 1）。NII が中心となり、超高速の通信ネットワークで臨床現場とクラウド基盤、さらに大学等の画像処理研究者をつなぐことで、大容量となる医療画像データ群を超高速で蓄積し迅速に研究活用する仕組みを整えました。このクラウド基盤には、CT 画像に限ってもすでに 1 億 6 千万枚以上の画像が格納されています。また、全国の様々な臨床機関から収集されたデータベースとなっているのも特色です。これらの仕組みにより、本クラウド基盤と SINET5 を活用した医療画像 AI 研究が可能となり、全国の画像処理研究者が集結しています<sup>(\*)9)</sup>。

このクラウド基盤は、COVID-19 肺炎に限らない臨床データが悉皆的・網羅的に収集されています。今回、この膨大なデータから、COVID-19 肺炎の AI 解析用プラットフォームを迅速に整備できたこととなります。つまり、今回の NII チームの成果は、ターゲット疾患を定めずに収集した過去のデータから、状況に応じて該当するデータを自動選別し、目的に最適な AI 機械学習用データセットを極めて迅速に整備できる技術を開発したことを意味します。このことは、このクラウド基盤や超高速ネットワーク SINET5 を将来も継続して整備・運用すれば、未知の感染症などの国家的な緊急課題への対処にも有効だと確かめられたといえます。

## 【研究プロジェクトについて】

本研究成果は、日本医療研究開発機構（AMED）「臨床研究等 ICT 基盤構築・人工知能実装研究事業」「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」の助成を受けたものです。

〈メディアの皆様からのお問い合わせ先〉

**大学共同利用機関法人 情報・システム研究機構 国立情報学研究所**

総務部企画課 広報チーム

TEL:03-4212-2164 E-mail : [media@nii.ac.jp](mailto:media@nii.ac.jp)

〈AMED 事業に関すること〉

**国立研究開発法人日本医療研究開発機構**

ゲノム・データ基盤事業部健康・医療データ研究開発課

E-mail : [rinsho-ict-ai@amed.go.jp](mailto:rinsho-ict-ai@amed.go.jp)

- 
- (\*1) COVID-19 肺炎典型度：CT 画像を観察し典型的な COVID-19 肺炎に見えるものから肺炎とは解釈できないものまで 4 段階に分類したものを。放射線診断専門医が CT 画像を以下の 4 つに分類した。1. COVID-19 肺炎に典型的な所見を有するもの、2. 典型的とは言えない非特異的な所見で不確定なもの、3. 非典型的なもの、4. 肺炎の所見がないもの。これらを 2 群に分けて、COVID-19 肺炎典型度の高い疑わしい症例（1と2）とそうではない症例（3と4）とした。
  - (\*2) 小田昌宏, 林雄一郎, 大竹義人, 橋本正弘, 明石敏昭, 森健策、「COVID-19 症例の定量評価のための CT 像からの肺野自動セグメンテーション」、第 39 回日本医用画像工学会大会、2020 年 9 月 17 日（木）
  - (\*3) Coronavirus Update (2020 年 9 月 10 日アクセス)  
<https://www.worldometers.info/coronavirus/>
  - (\*4) 厚生労働省 国内の発生状況など (2020 年 9 月 10 日アクセス)  
<https://www.mhlw.go.jp/stf/covid-19/kokunainohasseijoukyou.html>
  - (\*5) Simpson S, Kay FU, Abbara S, et al.: Radiological Society of North America Expert Consensus Statement 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
  - (\*6) Ai T, Yang Z, Hou H, 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
  - (\*7) 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：  
詳しくは <https://www.nii.ac.jp/research/projects/amed/> を参照。
  - (\*8) SINET5 : Science Information NETWORK 5。詳しくは <https://www.sinet.ad.jp/> を参照。本研究では、通信相手を特定したプライベートな専用回線であるかのように仮想的に利用するサービスである VPN (Virtual Private Network : 仮想閉域ネットワーク) により医療画像等を転送している。
  - (\*9) Murao K, Ninomiya Y, Han C, Aida K and Satoh S: Cloud platform for deep learning-based CAD via collaboration between Japanese medical societies and institutes of informatics. Proc. SPIE 11318, Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications, 113180T (2 March 2020) <https://doi.org/10.1117/12.2543521>