

# 3. 医療ビッグデータ研究センター における AI 研究開発の取り組み

佐藤 真一 / 合田 憲人 / 村尾 晃平 / 二宮洋一郎

国立情報学研究所医療ビッグデータ研究センター

医用画像に対する深層学習を用いた研究は国内外で盛んに進められており、その診断精度は専門医のレベルに迫っているとの論文も発表されている<sup>1), 2)</sup>。しかしながら、実用化を目の前にして、学習データの量・質は十分なのか、日々発展を見せる画像処理の技術に対して十分な精度・性能が得られているのかなど、疑問や不安な点がある。特に、深層学習の精度・性能はデータの質や量に依存するので、大量のデータが必要となる。しかしながら、医療データの場合、個人情報から派生したデータであるという特殊性により、データ収集は施設や条件が限定されたものになってしまう。

そこで、国立研究開発法人日本研究開発機構（以下、AMED）の支援の下、「臨床研究等 ICT 基盤構築・人工知能実装研究事業」（以下、プロジェクト）が 2017 年から始まった。また、それを契機に同年 11 月には、国立情報学研究所（以下、NII）に医療ビッグデータ研究センターが設立された。このプロジェクトでは、「わが国の医療の質向上・均てん化・診療支援、及び日本発の医療技術の臨床開発に必要なエビデンスを提供するため、臨床研究等にかかる ICT に関する先端的技術を有する機関等を中心に、臨床研究等の ICT 基盤構築に関する研究を推進する」こととなっている。そのために、NII の保有するクラウド基盤構築のノウハウやネットワーク環境、全国の大学・研究機関の研究者との人脈を活用して活動を行っているので、その状況を紹介する。

## プロジェクトのねらい

深層学習に基づく人工知能（AI）画像診断技術の開発のためには、対象となるデータの量と質の確保が重要である。したがって、複数の施設からの<sup>しつぱい</sup>データ提供が必要である。また、扱うデータが医用画像であるため、たとえ匿名化されていても、強固なセキュリティが求められる。画像解析については、そのセキュアな環境の中で処理できることが必要である。これらの要求を満たす基盤構築を実現することが第一のねらいである。

画像解析については、国内の情報系研究者で医用画像に意欲的に取り組んでいる研究者が多数いる。日々発展している画像解析の手法や評価の仕方を取り入れられるよう、これらの人材を集め、タスクを共有して取り組む環境を構築するのが第二のねらいである。

また、上記の基盤と環境の上で、解析研究の結果を医師と共有し、臨床的に意義あるモノや仕組みを描くことも第三のねらいとなっている。

## プロジェクトの体制

プロジェクトでは、第一のねらいを実現、複数施設から悉皆的にデータを集めるために、画像種別ごとに専門医を擁する学会が集約する体制とした。これらの学会は画像解析のタスク設定も行う。医師個人の興味や得意分野に偏ることなく、医師からのニーズの高いタス

クを選定できることが期待できる。現在 6 つの学会が参加しており、その内訳は、2017 年度に AMED プロジェクトに採択された日本消化器内視鏡学会、日本病理学会、日本医学放射線学会、日本眼科学会、2018 年度に採択された日本超音波医学会・日本皮膚科学会である。データ解析のためのセキュアな環境については後述する。

前項の第二のねらいに対応して、国内の解析研究者を結集するために、国内の画像解析を得意とする情報系の研究室と協力して研究を進めている。2019 年 4 月現在で参加している組織は、東京大学、名古屋大学、九州大学、奈良先端科学技術大学院大学、中京大学である。今後、扱うデータの種類やタスクの増大に合わせて、さらに別の組織の研究者に参画していただくことも計画している。プロジェクトの体制を図 1 にまとめる。

## クラウド基盤の詳細

クラウド基盤全体は、セキュリティを確保するためインターネットには接続されず、国立情報学研究所が運用する学術情報ネットワーク（SINET）から提供される L2 VPN 接続のみとなっている。

### 1. データの収集について

画像データの発生源は各学会に関連する大学や病院であり、各関連学会がそれらのデータを収集する。各学会から NII のクラウド基盤に送信するデータに