

厚生労働科学研究費補助金

政策科学総合研究事業（臨床研究等 ICT 基盤構築・人工知能実装研究事業）

分担研究報告書

クラウド上の医療 AI 利用促進のためのネットワークセキュリティ構成類型化と
実証及び施策の提言

研究代表者 岡村 浩司 国立成育医療研究センター 室長

研究分担者 松井 俊大 国立成育医療研究センター 医員

研究要旨

ディープラーニングを中心とした人工知能の発展は医療にも大きな影響を与え、我が国では内閣府が主導する戦略的イノベーション創造プログラム等により、医療の質や効率の向上だけでなく、地域格差の解消、医療従事者の負担軽減など多岐にわたる検証が進められている。国立成育医療研究センターでは 2018 年の AI ホスピタル事業採択を機に、研究所の研究者、病院の医師だけでなく、全ての職員を対象としたデータサイエンスの啓蒙、教育活動からスタートし、顕微鏡写真からの感染症起因菌同定支援、病理画像のグレーディング支援システムなどの開発を行い、実際の診療における補助手段としての試用も開始した。クラウドシフトによる大小さまざまな医療機関からの利用を見据え、医療 AI プラットフォーム技術研究組合との協力体制でこれらウェブアプリケーションのコンテナ化と仮想デスクトップ基盤を介した安全な通信環境の構築も行い、最終的な目標は、各医療機関の電子カルテ端末からこのようなサービスを自由に、かつ安全に利用できるようにすることである。しかしながら、ランサムウェアをはじめとするサイバー攻撃の危険性が高まり、個人情報保護や患者不利益等への配慮がますます求められる状況にあっては、従来からの境界型防御に加え、ゼロトラストを前提としたシステム構築を避けることはできない。それだけでなく、信頼される安全性の実証が不可欠であり、かつ現状の最大の課題ともなっている。安心できる環境の実現は、患者および市民参画をも促し、ビッグデータに依存する医療 AI のさらなる発展を期待することもできる。本研究では、感染症起因菌同定支援サービスの開発、クラウドからの公開、そしてユーザが安全かつ信頼性の高い環境下で利用できるように SASE アーキテクチャの試験的な導入、さらにはセキュアなブラウザ経由での利用を想定したインターネット分離を実装し、電子カルテ端末から医療 AI サービスを安全に、そして安心してアクセスできる環境を整えるための技術検討を行い、現状の課題を明確にした。

A. 研究目的

ディープラーニングによる画像認識の飛躍的な精度向上はその後の社会を大きく変えることとなった。皮膚がんの診断など医療における AI の有用性が示されて以来、医療の質や効率の向上、地域格差の解消、医療従事者の負担軽減などを目指した医療 AI の研究開発が盛んに進められている。自動車の自動運転や、世界最強棋士を破った囲碁プログラムで注目を浴びた強化学習についても、個別医療の最適化、手術に使われる医療ロボットの制御など、さまざまな活用が考えられている。これらの技術はハードウェアとソフトウェアの技術開発をも促進してきた。その結果、さまざまな実行環境がクラウドとオンプレミスで構築された複雑なシステムの上に組み合わされている状況を作り出し、一方で個人情報保護や患者不利益等への配慮が求められる時代背景にあって、ランサムウェアをはじめとするサイバー攻撃の危険性がますます高まっている。

我が国では内閣府が主導する戦略的イノベーション創造プログラムにより AI ホスピタルの取り組みが 2018 年から始まり、国立成育医療研究センター(NCCHD)は、医療 AI プラットフォーム技術研究組合(HAIP)とともに採択され、医療データを共有し、一体となって医療 AI サービスの開発を進めてきた。国内多くの医療機関がこのようなサービスを、安全に、安価に利用できる環境を提供することを目的に、共同で調査等も行っている。今回、NCCHD で開発を進めていた感染症起因菌同定支援のサービスをコンテナ化し、HAIP の医療 AI プラットフォームから公開することを試みた。

独自サービスを用意することで、安全な利用環境の検討にさまざまな最先端技術を試すこともできる。本研究では、環境をサーバ

上に集約させる仮想デスクトップ基盤(VDI)、ゼロトラストの一ソリューションである SASE、またセキュアブラウザを利用するインターネット分離を取り上げ、いずれ実現させたい電子カルテ端末からの医療 AI サービス利用を検討した。

B. 研究方法

感染症起因菌同定支援においては、NCCHD における小児菌血症患者の検体から 23,753 画像の顕微鏡写真データを取得し、Microsoft VoTT を利用して 347,234 箇所を手作業で切り出し、16 種類の細菌真菌を区別するアノテーションを行った。ラベルと位置情報を JSON 形式で出力し、画像分類のための切り出しや、物体検出のための YOLO 形式への出力は Python スクリプトを用いた。

最初に TensorFlow を利用して画像の分類を試みた。データ拡張には Keras を利用した。ImageNet で訓練された Inception V3 の転移学習を Hitachi SR24000/DL1 を用いて実行した。物体検出については、Microsoft Azure コンピューティング インスタンスのサイズ Standard_NC12s_v3 を利用して構築された HAIP の AI 開発基盤を利用した。アルゴリズムは、PyTorch を基盤とする YOLOv5 を採用した。CentOS 7 に設定した YOLOv5 を用いて訓練を行い、起因菌の物体検出モデルを作成した。

コンテナ作成は、CentOS 7 に設定した Docker にて行い、デプロイ確認は Minikube を利用した。ウェブアプリケーションとしての公開は Amazon ECS を利用し、また、HAIP サービス事業基盤からの公開は Azure Kubernetes Service を利用した。VDI クライアントは Microsoft Remote Desktop、サーバは Azure Virtual Desktop でクラウドの Windows にアクセスし、そのブラウザから

HAIP サービス事業基盤にデプロイされているコンテナにアクセスさせた。ウェブカメラの制御は JavaScript のメディアストリーム API に含まれている `getUserMedia()` を利用し、クラウドの Windows に接続されているデバイスを動作させた。

C. 研究結果

小児科として高度先進医療を提供する NCCHD は、免疫不全、臓器移植後に免疫抑制剤の投与を受けているなど、細菌感染に関してハイリスクな患者を多数抱えている。適切な抗生物質の選択など治療方針の決定は患者の予後に直結し、また不必要な抗生物質の使用は耐性菌の問題もあり、責任ある対応が求められている。そこで菌血症患者の検体の顕微鏡写真から起因菌を迅速に同定する医療 AI の開発を行なった。

小児の菌血症患者の血液培養からグラム染色を行なって得られた顕微鏡写真に、生化学反応や質量分析によって決定された細菌や真菌の種類を正解ラベルとして教師データを作成した。まず、10 μm 四方のクroppに対してアノテーションを行い、グラム陽性桿菌、グラム陰性桿菌、グラム陽性球菌、グラム陰性球菌、それから背景の5分類を試みた。Inception V3 の転移学習により訓練を行ったモデルに対し、ランダムに切り出した10 μm 四方のクroppでテストを行った。多数のクroppに対する結果のうち、背景を除いた多数決を取ることで良好な結果が得られることを確認した。

感染症起因菌をより詳しく同定するため、15 細菌および1 真菌を物体検出で区別するアノテーションを行なった。内訳は、腸内細菌科、緑膿菌、エンテロコッカス属、コアグラエ陰性ブドウ球菌、バシラス属、黄色ブドウ球菌、B 群 β 溶血性レンサ球菌、レン

サ球菌属、ヘモフィルス属、肺炎レンサ球菌、化膿レンサ球菌、リステリア、シュードモナス属、コリネバクテリウム、グラム陰性球菌、カンジダであり、これらに赤血球を加えた17ラベル、23,753 画像から347,234 クroppの切り出しで訓練を行なった。モデルはYOLOv5xを採用し、V100 を搭載するHAIPのAI開発基盤で、150 エポック34時間をかけて独自モデルを完成させた(図1)。

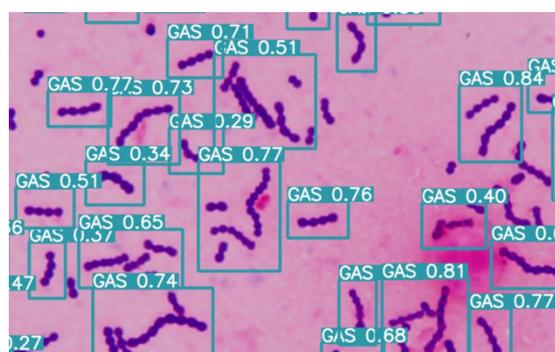


図1.化膿レンサ球菌の検出

検出モデルをウェブサービスに組み込み、ブラウザからの画像データアップロードで瞬時に、かつインターネット接続ができればどこからでも結果が得られる仕組みを作り上げた。そして Docker によりコンテナ化を行い、作成したコンテナがオンプレミスに加え、AWS でも動作することを確認し、扱う医療データを安全にやり取りする目的で、HAIP サービス基盤が導入している Microsoft Azure の VDI 環境で公開することができた。利用可能ユーザは限定しているものの、申請があれば誰でも利用できる体制を整えた。

ユーザはリモートデスクトップクライアントを利用して、クラウドのブラウザにアクセスし、デスクトップ画像の通信で利用する形態である。顕微鏡写真はデータをアップロードすることもできるが、顕微鏡に接続され

たコンピュータからの簡易的な利用を見据え、ディスプレイに表示された画像をウェブカメラで撮影して検出できるように設計されている(図 2)。ユーザが手元で操作するウェブカメラからのリアルタイム入力は VDI により、直結されたクラウド側に存在するウェブカメラでのデータ扱いとなる。



図 2. VDI でクラウドのウェブカメラを利用

さらにクラウド環境のネットワークアクセスの安全性を確保するため、ゼロトラストの一ソリューションである SASE を導入することとし、実際には Cato Networks 社がサービスを展開している Cato SASE クラウドプラットフォーム (CATO) の調達を行い、利用できる環境が整った。またセキュアブラウザを利用するインターネット分離については、ジェイズ・テクノロジー社の RevoWorks

を調達し、セキュリティ対策の選択肢を増やした。このような状況で NCCHD の電子カルテネットワークでの試用を打診したが、残念ながら一時的な利用であっても許可は得られていない。

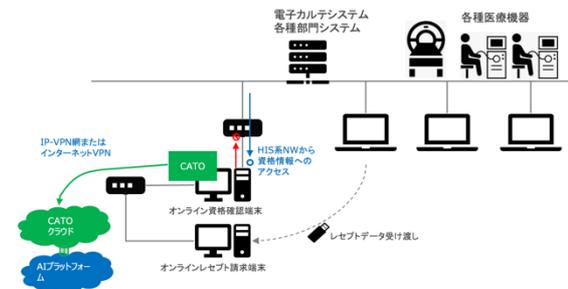


図 3. SASE として CATO の導入

D. 考察

感染症起因菌同定支援システムは臨床検査技師をはじめとする医療関係者の力量を上回る結果が出ている一方で、訓練データが単一機器から習得されていることもあり、他施設で期待したほどの認識精度が得られていないという実態がある。実装可能であることは示すことができたが、社会実装を進めるためには、より広範囲なビッグデータの収集が必要になる。

これら医療 AI サービスは開発段階ということもあるが、そうでなくとも使用頻度が高いとは言えず、クラウドの仮想マシンやコンテナとしての運用では、24 時間 365 日の連続稼働となり、ほとんど使われていないのに課金される状態が続くことになる。このような状況に対し、サーバレスアーキテクチャを採用した場合、必要最小限のリソースで、API のコール数とデータ転送量に対してのみ課金が発生するため、コストを大幅に下げることができる可能性がある。サーバのセットアップ、メンテナンス、スケーリングといったインフラ管理からも解放される。各コンポ

ーネットには自動的に復旧される仕組みも備わり、耐障害性の高い仕組みを作り上げることができはずで、さらにはアイデンティティベースの認証など、ゼロトラストの考え方と一致する面もある。次年度の取り組みとして検討を進める必要があると考えている。

E. 結論

クラウド上の医療 AI サービス利用促進のためのネットワークセキュリティ構成をゼロトラストの観点も含めて提案し、電子カルテネットワークあるいは電子カルテ端末からの利用を試したいと考えていたが、境界型防御を通してきた電子カルテのネットワークを改変することは、技術面だけでなく、セキュリティに対する不安や、管理者の責任問題から非常に困難である現状が明らかとなった。ゼロトラストを謳う統合的なサービスの仕様を眺めると、複雑で分かりにくく、必要性が不明な高機能、それに付随する高価格を目にすれば、気安く導入できるようなもの

ではない。また、ゼロトラストの言葉をもって管理者、さらには患者や市民に安心を与えられる状況ではない。ゼロトラスト・セキュリティは開発者側を鼓舞する上で便利な言葉ではあるものの、ユーザ側を納得させ、医療 AI サービスのより広範な活用を図るためには、セキュリティ対策の実績を積むとともに、分かりやすい説明を続けるという地道な作業が必要であるように思われる。

F. 健康危惧情報

総括研究報告書に記載

G. 研究発表

岡村 浩司, 松井 俊大. 電子カルテ端末からの利用を見据えた医療 AI サービスの開発. *医療情報学*, 2024, **44**(Suppl.), 354-357

H. 知的財産権の出願

なし

