



数少ない網膜剥離画像から高精度AIを作るまで—眼科医・升本浩紀が語る「医療AI応用までの道のり」(4)

2019年11月8日(金)

ツカザキ病院眼科で眼科医兼人工知能エンジニアのチーフとして働く升本浩紀氏が、自らのAI開発の経験をベースに、医療AIの現状について紹介する連載コラムです。

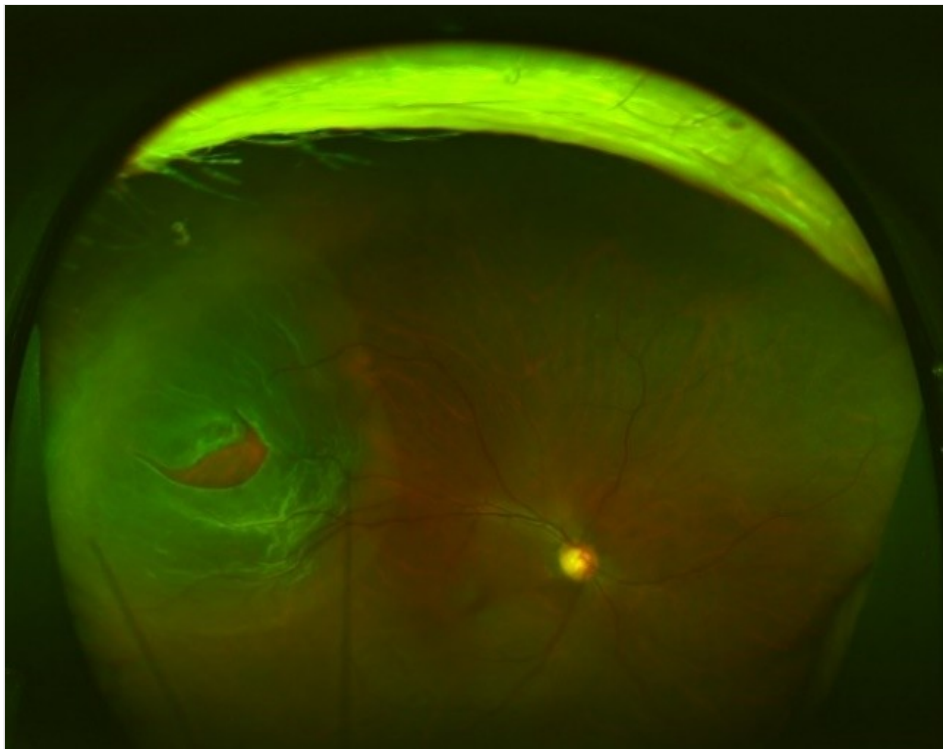
» [連載1回目から読む](#)

本邦における医療AIの開発を進めるにあたって避けては通れない問題が、個人情報保護によって十分な医療用画像を集められないという点です。日本がいくらAIの技術を向上させて、「数少ない枚数で学習させられる方法を構築しました!」という研究を行っても、やはり十分な数を用いて学習させたAIに勝つことは至難の業だというのは厳然たる事実として存在します。個人情報保護の制限がゆるい国に、医療AIの開発で勝つためにはどうしたらいいのでしょうか。

医療資源が豊富で、少なくとも現時点ではAIがなくても一切困っていない日本の医療現場において、プライバシーや個人情報保護への価値観が変わるとはとても思えません。データ量の問題を乗り越える方法が必要不可欠です。

不足する網膜剥離の眼底画像

例えば、我々はこれまで、「超広角眼底撮影装置 (Optos200Tx)」で撮影した画像を用いた各種疾患のAI診断論文を発表してきました。その中でも、網膜剥離の診断には非常に重きを置いています。その理由は、米国では網膜剥離が眼科関連における訴訟関連疾患の第一位であるためです。網膜剥離は早期に治療を適切に行えば、ほぼ視機能は保たれる一方で、治療が遅れると失明につながる病気だからです。つまり、網膜剥離を見逃すことは失明につながるリスクを非常に高めるということであり、結果的に訴訟につながるのです。そこで我々は失明を防ぐために、網膜剥離を見落とさないためのAIの構築を非常に重要視しています。



網膜剥離のOptos画像

しかし、私が所属するツカザキ病院眼科は年間手術件数が1万件を超えるほどの日本最大規模の眼科臨床施設であるにもかかわらず、約1000枚しか網膜剥離のOptos画像を所持していません。網膜剥離は1万人に1人ほどの割合といわれており、糖尿病性網膜症や緑内障、加齢黄斑変性と比較してまれな病気だからです。

一方で、例えば糖尿病性網膜症においては3万枚ほどのデータセットが無料で公開されています。(参考URL:<https://www.kaggle.com/c/diabetic-retinopathy-detection/>) 1000枚という枚数は一般的なAI研究において、少ない枚数だと言わざるを得ません。

少ない枚数でAIを学習させるためには？

では、枚数を増やすにはどのようにしたらよいのでしょうか？ 最も簡単な方法は、データを多施設で共有する方法です。しかし、先述のように個人情報の取り扱い問題が叫ばれる中、リスクを取って共有したがる施設は非常に稀です。特に、最も研究に力を入れている大学病院などは規制が多く、仮にデータの共有をしようとしても、実際にデータの共有がなされる頃には研究の旬が過ぎているのは明らかです。

次善の策として考えられるのは、データそのものを共有するのではなく、モデルを共有するという手法です。これを「分散協調学習 (Federated Learning)」と呼びます。

それぞれの分院で作ったAIを「統合」する

例えば、中央病院がその疾患の画像をわずかしか持っていない一方で、複数の分院がある程度の枚数を持っているという状況を思い浮かべてみてください。その際に、

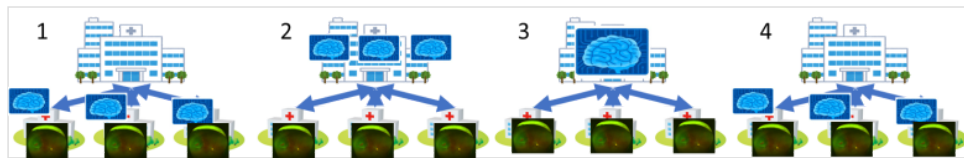
(1) 各分院の内部のサーバーなどで、それぞれにモデルを育成する。

(2) 各分院で育てられたモデルを中央病院に転送する。

(3) それらを統合(平均化)する。

(4) 統合されたモデルをそれぞれの分院に再配布し、さらにその分院の中で各分院の画像をもとにモデルを育成していく。(2に戻る)

このようなプロセスを繰り返すことで、一つの病院内での画像から得られた知識だけではなく、他の病院の画像から得られた知識も吸収していくことができるようになります。それも画像データそのものを一切外部に出すことなくです。これが、「分散協調学習 (Federated Learning)」です。



我々はこの分散協調学習を用いて、検査大国であり画像などの検査データはたくさんあるにもかかわらず、プライバシー意識が高くAI開発を進めづらい現状を打破しようと考えています。

画像情報の転送無しに精度高いAIを作成するために

前回の連載で、AI開発を行うにあたっては、簡単なトライアルをまずは行う必要があるということをお述べてきました。そのため、我々は院内のデータのみですが、それを別々に扱うことで「疑似」分散協調学習の検証を行いました。

具体的にはまず、ツカザキ病院眼科データベースから抽出した正常眼底画像973枚および網膜剥離画像973枚を検証に使用しました。これを、訓練データが778枚、テストデータが195枚になるようにそれぞれ分割しました。その上で、ここでは下記2つの条件をもとにAIを作成し、テストデータにおける性能を評価しました。

【条件1: Normal model】訓練用データを全て一気に学習させ、モデルを作成する

【条件2: Federated model】訓練データを3分割し、それぞれ分割した画像群をバラバラに使用し、最後に統合することで、モデルを作成する

条件2の工程について詳細に説明します。

1) まず、訓練用データである正常眼底画像 778枚および網膜剥離画像 778枚を3分割します。



2) これらの各訓練用データをもとに、モデルを別々に訓練します。



3) 各モデルを統合(平均化)し、テストデータで検証します。



4) モデルを3つにコピーし、2)に戻ります。



その他のNetwork構造やハイパーパラメータなどはNormal modelと全く同一としました。その結果、Federated ModelはNormal modelをわずかに上回る結果となりました。

	Federated Model	Normal Model	p value
AUC	0.881 (95%CI: 0.846 – 0.915)	0.874 (95%CI: 0.838 – 0.909)	p=0.33
Sensitivity	80.5% (95%CI: 74.2% – 85.8%)	77.9% (95%CI: 71.5% – 83.6%)	p=0.62
Specificity	80.5% (95%CI: 74.2 – 85.8%)	78.5% (95%CI: 74.2 – 85.8%)	p=0.71

結果の解釈としては、学習するたびにパラメータを平均化したため、Federated modelの学習がNormal modelの学習に比べて簡単に収束した可能性が考えられます。今回の研究はあくまで疑似的な分散協調学習であり、データセットの数も少なく、検証の妥当性もそこまで高いとはいえません。

しかし、本研究の結果はデータそのものを転送することなく、十分な性能のAIを作成することができ的可能性を示唆しています。この手法はAI開発における機密性や個人情報の問題を克服できる可能性があると考えています。

連載4回目の今回は、AIを作る際の技術ではなく、社会的課題であるプライバシーやデータ倫理の問題をいかに乗り越えるかの方法の一つである分散協調学習を紹介しました。検査大国である一方、データの取り扱い規則が厳しい日本がAI開発競争で勝つためには、このような方法を使うしかないと個人的には思います。日本の医療がGoogleなどの外資系ソフトウェアに依存しないためにも、今後、分散協調学習を駆使することがさらに重要となっていくでしょう。



升本浩紀
ツカザキ病院 眼科 医師／株式会社シンクアウト 最高技術責任者

2016年 東京大学医学部卒業。在学中に中小企業診断士や公認会計士試験に合格。2018年からツカザキ病院。眼科医として臨床を行う傍ら、医療AIの研究・開発に取り組んでいる。日本眼科AIのトップランナーとして国内外の学会や、医師、医学生向けの講演を多く行っている。関心領域はオペレーションマネジメントや、スマートフォンを用いたビジネス。好きな人工知能フレームワークはPyTorch。

このシリーズの記事

- 

数少ない網膜剥離画像から高精度AIを作るまで—眼科医・升本浩紀が語る「医療AI...
2019年11月8日 (金)
- 

右眼と左眼を見分けるAIが実臨床で運用されるまで—眼科医・升本浩紀が語る「医療...
2019年10月9日 (水)
- 

「訴求力の高い」AIの構築に向けて—眼科医・升本浩紀が語る「医療AI応用まで...
2019年8月16日 (金)
- 

眼底画像からAIが診断を下すしくみ—眼科医・升本浩紀が語る「医療AI応用まで...
2019年7月11日 (木)

[シリーズ一覧](#)

関連カテゴリー

- 画像診断支援
- 解説・コラム
- 眼科
- 眼科医・升本浩紀が語る「医療AI応用までの道のり」

参加募集中のAIラボプロジェクト

- 

医療×AIセミナーシリーズ第10回
「循環器とAI」(2019年11月30日
開催)
M3 Supported 2019年10月24日 (木)
- 

医療×AIセミナーシリーズ第9回
「AI時代の医療とトラスト:日仏哲
学対話」(2...
M3 Supported 2019年8月22日 (木)
- 

【締め切りました】医療×AIセミナ
ーシリーズ第8回「医療者がゼロ
から学ぶ、AI...
M3 Supported 2019年7月25日 (木)
- 

【締め切りました】医療×AIセミナ
ーシリーズ第7回「臨床現場を効
率化するAI・...
M3 Supported 2019年6月20日 (木)

[AIラボ プロジェクト一覧](#)

関連する AIラボニュース

- 

画像データベース構築や機器承
認で日本のAI診断はどう変わる
か—第73回日本臨床...
2019年11月8日 (金)
- 

AI診断で真っ先に改革が進む眼
科医療で備えるべきことは—第
73回日本臨床眼科...
2019年11月6日 (水)
- 

気胸検出AIシステムがFDA認可
—轟・加藤の「医療AI トレンドを
追う」(9)
2019年10月25日 (金)
- 

現実化した「超個別化医療」高額
費用は誰が払うべきか？
2019年10月24日 (木)

[AIラボニュース一覧](#)

開業を検討中の先生

開業意向に関するアンケートに回答して、最大🎁10p進呈

回答する >>

