共同企画 | 第40回医療情報学連合大会（第21回日本医療情報学会学術大会） | 共同企画
共同企画8
医学医療における AI応用
2020年11月20日(金) 09:00 〜 11:00 B会場 (コングレスセンター3階・31会議室)

[3-B-1-02] 医用画像診断における深層学習モデルの開発

*小林 和馬[1]（1. 国立がん研究センター研究所）
*Kazuma Kobayashi[1]（1. 国立がん研究センター研究所）

キーワード：Deep learning, Medical images, Structured database

コンピュータの計算能力の加速的な向上と深層学習に代表されるアルゴリズム革命によって、特に画像認識の一部の課題に対しては、機械による識別性能が人間の識別限界を超えるに至ったと言われている。 OECD諸国の中で、 CTや MRIのなどの放射線画像診断装置を人口あたり最も多く有している本邦は、膨大な医用画像を蓄積している。こうした医用画像は、とりわけ現代のがん診療において、疾患の診断、治療方針の決定、治療及び経過観察のいずれの過程においても大きな役割を果たしている一方で、データ分析の視点から見ると、不均質で構造化されていない。すなわち、日々の臨床から生み出される質的、量的に膨大なデジタルデータを、良質で構造化された大規模なデータセットへと変換し、最先端の人工知能技術を用いた価値創造の契機を創出することが重要である。　こうした視点に立って、国立がん研究センター中央病院・研究所では、人工知能技術を活用した統合的ながん医療システムの構築に取り組んできた。本講演においては、病院内の医用画像をコンピュータによる解析、学習、推論の対象となるようなビッグデータへと効率的に変換するための我々の取り組みについて紹介する。特に、医用画像を起点とし、診療情報を含めて構造化するためのプラットフォームについて、我々のプロトタイプを取り上げる。続いて、一般画像と比較した際に医用画像において考慮するべき特殊性について問題点を整理し、それぞれについての技術的な解決の方向性に関して具体的な事例と共に発表する。特に、臨床医が医用画像を読影する際に、異常な部位に本来存在するべき正常な解剖学的情報を想像し、そこからの画像情報の差分を頭の中で評価していると言える。こうした臨床医の認識過程を再現するような医用画像のアルゴリズム構築に取り組んでおり、その一部について紹介し、有効性について議論させて頂きたい。

©一般社団法人 日本医療情報学会

医用画像診断における深層学習モデルの開発

小林和馬[*1]

*1 国立がん研究センター研究所がん分子修飾制御学分野

Applied Deep Learning in Medical Image Diagnosis

Kazuma Kobayashi[*1,2]

*1 Division of Molecular Modification and Cancer Biology, National Cancer Center Research Institute

Medical images play an important role in the processes of diagnosis, treatment, and follow-up. However, a vast amount of medical images accumulated in hospitals are heterogeneous and unstructured from the viewpoint of data analysis. Therefore, it is important to transform the digital data generated in daily clinical practice into a high-quality, structured data so that they can create opportunities for the development of machine learning models. From this perspective, we have been working on the construction of an integrated cancer treatment system driven by artificial intelligence technology. In this talk, we will introduce our efforts to build a system to effectively transform medical images in hospital into big data for computer analysis. Further, we will summarize the problems that need to be considered for the development of deep learning models using medical images in comparison with general images. Especially, it can be said that clinicians imagine the normal anatomical information that should have existed in the abnormal area and evaluate the difference of local features from it in their mind when they diagnose the medical images. We have been working on developing algorithms for medical images that can reproduce the recognition process of clinicians.

Keywords: Deep learning, Medical images, Structured database.

1. はじめに

CT や MRI などに代表される医用画像は、現代の医療において、疾患の診断、治療方針の決定、治療および経過観察のいずれの過程においても大きな役割を果たしている。また、加速的に高度化するデジタル技術に伴って、医用画像が内包する情報も、人体の解剖学的構造のみならず、臓器の機能や疾患の悪性度といった豊富な定量化情報へと進歩してきた。その結果、豊富な画像モダリティから生み出される、質的、量的に膨大なデジタルデータが日々の診療において蓄積されてきた。

人工知能の代名詞となった深層学習を始めとする様々な機械学習アルゴリズムを用いて、病院に蓄積される医用画像を貴重なデータ資源として活用し、臨床現場における様々な課題解決を目的とした研究開発競争が世界的に激化している。特に、本邦はOECD諸国の中でも、人口当たりで最も多くの画像診断装置を保有している。超高齢化社会を迎えていよいよ逼迫する本邦の医療の持続可能性という視座からも、実臨床において有用な人工知能技術を確立することは急務の課題である。

本稿においては、人工知能技術を用いた臨床支援システムを確立するために、良質で構造化されたデータセットを効率的に収集するためのプラットフォームの開発を目指した国立がん研究センターにおける取り組みについて紹介する。続いて、既に概念実証から社会実装・製品開発の段階に至ったと考えられる深層学習技術を、医療現場に実装するに当たっての実際的な課題について議論したい。更に、臨床医は異常所見を観察する際に、そこに本来あるべき正常所見を想起し、そこからの差分を評価していると言える。こうした医師の認識過程を再現するための画像処理アルゴリズムの開発についても紹介する。

2. 臨床データ構造化のためのアノテーション・プラットフォームの提案

深層学習を始めとする様々な機械学習アルゴリズムを学習するためのアプローチは、教師あり学習と教師なし学習に大別することができる。教師あり学習では、入力データに対して正解ラベルが対になるように与えられたデータセットを用意し、学習機械に入力データが与えられた際に、その出力が正解ラベルに近づくようにパラメータの調整が行われる。一方、教師なし学習では、入力データに対になるべき正解ラベルが与えられない為、学習の目的は、入力データそのもののパターンや分布を推定することとなる。

一般に、特定のタスクに対する高精度な機械学習モデルを実現するためには、大量のデータに基づいた教師あり学習が有用である。このためには、入力データに対して正解ラベルを付与する必要があるが、このラベル付けをアノテーションという。アノテーションするべき情報をどのように定義するべきかは、どのようなタスクを解きたいのかという解析の目的に応じて決定される。

画像認識のタスクは、三段階に区別して考えられることが多い。具体的には、「その画像が何なのか」を分類する画像分類、「その画像のどこに何があるのか」を検出する画像検出、「個々のピクセルが何を意味しているのか」を識別するセグメンテーションである。医用画像におけるアノテーションも、こうしたタスクの粒度を考慮する必要がある。すなわち、画像分類に対応するためには医用画像の特定シリーズあるいは特定スライスに対する分類ラベル、画像検出に対応するためにはバウンディング・ボックスなどで指示された病変部位の位置情報、セグメンテーションに対応するためにはピクセルあるいはボクセル単位で定義可能な輪郭情報などを付与することを要する。

また、医用画像の読影時に、医師がどのように画像の他の臨床情報を参照しているのか、という過程に対する洞察も重要である。医用画像における特定の所見を観察する際に、病変が存在する位置や広がり、性状などの画像的特徴の認識

に加えて、医師は、そこから推定される鑑別疾患に関連する他の検査や患者背景といった情報を併せて参照する。すなわち、機械学習モデルにより質の高い臨床診断支援を実現するためには、このような医師による情報の参照過程を可能な限り再現できるような、医用画像に紐づく形で格納された臨床情報の構造化データベースが必要である。

これを実現するために、医用画像対して、階層的に構造化されたアノテーション情報の必要性を提言する。

我々は、具体的に次のような 3 つの階層をもったアノテーションの定義を用いることで、医用画像を起点としたがん診療データの構造化を行っている。

患者単位:臨床経過中に変化しないラベル情報。がん診療における主要疾患名、TNM 分類、治療情報、転帰など。

画像シリーズ単位:画像検査単位で付与されるべきラベル情報。検査目的、検査時現症、あるいは領域性をもって定義しにくい画像所見(びまん性変化、構造変化の有無)など。

関心領域単位:領域性をもって定義されるラベル情報(原発腫瘍、リンパ節腫大など)。ここの関心領域は、領域についての情報とともにセマンティックな分類情報(原発腫瘍の内部や辺縁の正常についての分類など)を保持してもよい。

しかしながら、一般的にアノテーション情報を細かくすればするほど、これを付与するための作業負担が増大する。また、こうしたアノテーションの作業と診療行為が、その内容として大きく乖離していることは、診療データ構造化のための大きな障壁であると言える。そこで、人工知能技術をうまく活用することによって、臨床業務に対する支援とできる限り両輪となった、診療データを構造化するためのプラットフォームを開発するべきであると考えた。本講演では、我々が開発した臨床データ構造化のためのアノテーション・プラットフォームのインターフェイスを紹介する。

3. 深層畳み込みニューラルネットワークと医用画像の特殊性

人工知能技術と総称され、実社会の幅広い領域において大きな影響を及ぼしつつある技術革新の中心に位置するのが、深層学習である。深層学習とは、深層ニューラルネットワークの学習法の総称である。深層ニューラルネットワークは、生物の脳神経系に着想を得た情報処理システムである人工ニューラルネットワークを多層化したものである。多層構造の人工ニューラルネットワークに関する先駆的な概念や手法は以前から提案されていたが、とりわけ近年になって、コンピュータの処理能力の飛躍的な向上、アルゴリズムの洗練化、ビッグデータの社会の各所での蓄積などと相まって、大きく成功するに至った。

特に、深層学習を従来の機械学習と比較して特徴付ける性質として、一貫学習(end-to-end learning)が挙げられる。すなわち、深層学習では、生の観測データから、人の手による特徴量の設計を行うことなく、低次から高次までの各段階の階層でタスクに応じた特徴抽出を行うことができる。優れた予測アルゴリズムはデータの性質を十分に捉えた良い情報表現に依存していることが多い。つまり、深層学習が自然言語、音声、画像などの幅広いモダリティにおいて高い認識性能を示す理由は、その基本部分をなす汎用的な特徴抽出アルゴリズムにあるといえる。

3.1 深層畳み込みニューラルネットワークの概観

医用画像を含む画像認識のタスクにおいて、広く使われる深層ニューラルネットワークの種類の一つが深層畳み込みニューラルネットワークである。深層畳み込みニューラルネットワークの最も典型的なものは、畳み込み層とプーリング層の交互の重ね合わせからなる。畳み込み層では、入力画像に対する畳み込み演算と活性化関数による非線形変換を行って、特徴マップと呼ばれる画像特徴の集合を出力する。プーリング層では、特徴マップにおける局所的なパターンを統合し、情報を抽象化して次の層に伝搬される。

深層畳み込みニューラルネットワークにおいては、その最も基本的な演算単位である畳み込みが重要である。畳み込みにおいては、画像よりも小さなサイズを持つフィルタと呼ばれる行列を用意し、画像の各画素に対してフィルタの中心を重ね、重なった領域全体に渡って画素とフィルタの内積を計算する。フィルタを適切に設計することによって、例えば、ぼかし、先鋭化、エッジ強調などの基本的な画像演算を実現することができる。実際、深層畳み込みニューラルネットワークの学習においては、逆誤差伝搬法と言われるアルゴリズムを用いて、各畳み込み層のフィルタの値が最適化される。興味深いことに、大規模な一般画像を用いて深層畳み込みニューラルネットワークを訓練することで、哺乳類の一次視覚野における特徴抽出機構をモデル化した Gabor フィルタに類似した畳み込み層が得られることが知られている[1]。

3.2 医用画像における特殊性

深層畳み込みニューラルネットワークは、画像認識における幅広いタスクにおいて高い性能を示し、社会の様々な分野で活用され始めている。一方、ここでは、一般画像と比較した際に、医用画像において考慮するべき特殊性について、主に三つの観点から議論したい。

3.2.1 識別するべき所見の局所性

対象オブジェクトが被写体として画像の中心に大きく写っていることが多い一般画像と異なり、医用画像において検出されるべき病変は、正常解剖の中の一部の領域に限局した異常陰影であることがある。そのため、局所情報の変化に対する高い感応性が求められる。

3.2.2 小さなサンプルサイズ

多くの一般画像のデータセットが、数万からときには数百万の単位で画像枚数を保有しているのに対し、医用画像では、一つの研究において高々数百から数千の単位のサンプルしか得られないことがある。

3.2.3 同一クラス内の不均一性

臨床医学において、「異常」として出現する病変の表現型は多種多様であり、可能性としては無限のばらつきがあり得る。とりわけ、がんという疾患は、一つの腫瘍内部においても多数のサブクローンにより構成されており、「不均一性」がその本質であるとまで言われている。すなわち、一つの疾患ラベルの中にも見た目が大きく異なる画像が含まれていることが多く、その隠れた層別化を適切に処理することが重要である。

特に、クラス内の不均一性についてはデータセットのバイアスとなりうる因子であり、モデルの性能を報告する際においては、可能な範囲で詳細に記述するべきであると考える。すなわち、同一の疾患を対象にしているつもりであっても、当該疾

患の多様なサブクラスのバリエーションが学習用データセットと実装先の現実環境との間で乖離してしまうことによって、期待された性能が発揮されないようなことが起こりうる。本講演においては、医用画像の特殊性に注意した画像認識アルゴリズムを構築するための、我々の試行錯誤について紹介する。

4. 存在するべき「正常画像」を仮定した医用画像認識アルゴリズムの開発

医用画像における病変の画像的特徴は、疾患により引き起こされる病的な変化と、そこに元々存在している正常な解剖学的特徴が組み合わさったものであると言える。人体の正常な解剖学的特徴は共通しているため、臨床医は、異常な領域に注目した際に、そこに背景として存在しているだろう正常な解剖学的特徴を差し引いて、異常さのみを純粋に反映する画像的特徴を想起し、評価している。そのため、疾患が発症した前後の同一の領域の医用画像を比較することや、類似した病変を有する異なる患者間の医用画像を比較することが、画像診断において非常に重要となる。

こうした視点に立って、臨床医が医用画像を読影する際の認識過程を再現できるような画像認識アルゴリズムの開発に取り組んでいる。そのうちの一つが、ラベル情報を不要とする教師なし学習である。

放射線画像において「正常」とみなされる所見は、集団の平均値周囲に分布し、出現頻度が高い。一方で、「異常」とみなされる所見の現れ方は多種多様であり、それぞれのパターンの出現頻度は相対的に小さいと言える。すなわち、「正常」と「異常」のそれぞれに分類される画像パターンの出現頻度の間には、大きな不均衡が存在する。そこで、データセットの大部分を占める「正常」な画像パターンを深層ニューラルネットワークに学習させることによって、正常解剖を表現する潜在分布をモデルの内部に構築することが出来る。続いて、何らかの異常を含む画像をこのモデルに入力すると、その入力画像に最も近い正常画像に置き換わったものを出力として得ることができる。そこで、画像的な類似性を評価する何らかの関数を用いて、ピクセル単位での誤差を評価することにより、入力画像中の異常部位を可視化することに成功した[2)]。

教師なし学習はモデル開発におけるアノテーションのコストを大幅に低減することができ、教師あり学習を補完するものとして有望である。その他、本講演においては同様の視点から構築した他の応用例についても紹介したい。

参考文献

1) Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM 2017; 60(6) : 84-90.
2) Kobayashi K, Hataya R, Kurose Y, et al. Unsupervised Brain Abnormality Detection Using High Fidelity Image Reconstruction Networks. arXiv preprint arXiv: 2005.12573