

一般口演

一般口演28

機械学習・アルゴリズム・解析モデル

2017年11月23日(木) 12:45 ~ 14:15 E会場 (10F 会議室1003)

[4-E-2-OP28-2] 教師なし深層学習による読影中の注視点軌跡からの特徴量抽出の試み

寺下 貴美, 堤 翔子, 佐藤 充, 土井 邦雄, 小倉 敏裕（群馬県立県民健康科学大学）

【目的】近年、コンピュータ支援診断の発展が著しい。しかし、コンピュータ支援診断が確立しても、人間の読影技術の向上は必要である。種々の診断画像の読影法は解剖、病理および撮像学的な見地から、経験も踏まえ確立されている。しかし、実際に人間がどのように画像を認識しているかを評価するのは難しい。例えば、ROC曲線での評価はアウトカムの評価であり、プロセスは評価しておらず、教育的には不十分である。読影過程を評価する一つの方法に、注視点の解析があり、多くの研究が報告されているが、単に移動距離や注視時間を分析するのみで、読影中の行動を評価してはいない。そこで我々は読影プロセスの評価を目的に、教師なし深層学習を用いて、読影中の注視点軌跡における特徴量の分析を試みた。

【方法】対象画像は10枚のCT कोरोノグラフィ腸管内部展開像とした。読影は診療放射線技師8名で行った。まず、1画像当たり約30秒で順次読影を行い、注視点データを得た。注視点データは0.02秒毎にモニタのXY位置で記録される。次に注視点を30点ごとにプロットした画像を軌跡画像とし、それらを学習セット10,426枚とテストセット206枚に分割した。この学習セットに対してオートエンコーダで教師なし学習を行い、16の分類フィルタを作成した。この分類フィルタを使い、テストセットを分類した。分類結果について軌跡画像と同部位の原画像を表示し、どのような行動かを特定した。

【結果】分類フィルタを大別すると、画像を走査している行動と一部分を注意深く観察している行動が特定できた。また両者の原画像には、病変のある画像と病変のない画像が含まれていた。これはアウトカム評価で偽陰性とされてしまうことが、病変があっても気付かない「見落とし」と、病変を認識していながらも病変ではないと判断した「見逃し」のような読影のプロセスを表しているのではないかと示唆された。

教師なし深層学習による読影中の注視点軌跡からの特徴抽出の試み

寺下貴美^{*1}、堤翔子^{*1}、佐藤充^{*1}、土井邦雄^{*1}、小倉敏裕^{*1}

^{*1}群馬県立県民健康科学大学

A pilot study for feature extraction from gaze tracking point in medical image interpretation using an unsupervised deep learning

Takayoshi Terashita^{*1}, Shoko Tsutsumi^{*1}, Mitsuru Sato^{*1}, Kunio Doi^{*1}, Toshihiro Ogura^{*1}

^{*1} Gunma prefectural of health sciences

Abstract

It would be useful to improve medical image interpretation for human observers even if computer aided diagnosis may be established as diagnostic tool. The eye movement analysis would be useful to evaluate the reading process, but this method has not been established. Thus, it is necessary to evaluate the usefulness of eye movement analysis for medical image interpretation process. The features of reading process from eye movement patterns can be extracted by deep learning method. The purpose of this study is to propose new methodology for evaluating the medical image interpretation process. In this paper, we attempted to characterize eye movement patterns from gaze point data during image reading using deep learning method. The auto-encoder was adopted as deep learning method in this study. The results indicated that eye movements extracted by one of the weights of hidden layer would have corresponded to the behavior for carefully observing the interesting area like polyp in CTC image. It can be assumed that the reading process may be clarified by filtering based on these weights in further investigation. Furthermore, it may be applicable to the education for medical image interpretation.

Keywords: medical image interpretation; gaze tracking; unsupervised deep learning; auto-encoder; computed tomographic colonography

1. 緒論

近年、コンピュータ支援診断の発展が著しく、マンモグラフィなどではすでに、臨床現場で利用が進んでいる。しかし、コンピュータ支援診断が様々な診断画像で広く普及したとしても、人間の読影技術の向上は必要である。

種々の診断画像の読影法は解剖学、病理学、および画像工学、機器学、撮像技術学の見地から、さらに経験も踏まえて確立されている。一般的に読影能力を評価する方法として、ROC解析がある。ROC解析での評価は病変部を発見できたか、発見できなかったか、または病変でない部分を病変と誤認したかのような、アウトカムを評価する。例えば読影の指導として、「ちゃんと画像を見る!」「知識が不十分だ!」などの一方的な指摘は、「病変を認識できていたにもかかわらず判断を誤った」や「知識は十分であるにもかかわらず見ていなかった」という、個々の学習ニーズに対しては不適切である。これはプロセスに基づいた指摘ではないからであり、これまで初学者は数を重ねることだけでカバーしてきた。

読影のプロセスを評価するには読影中の行動を取得し、画像を認識できているか否かを把握する必要がある。しかし、実際に人間がどのように画像を認識しているかを知るのには難しい。読影中の行動を取得する一つの方法に、注視点を追跡する方法がある。読影中の注視点の動きについて、これまで多くの研究が報告されているが、単に移動距離や注視時間を集計するのみで、読影中の行動を評価しているものは少ない。読影プロセスを評価するには、注視点を追跡したデータから新たな手法によって画像認識の知見を得る必要がある。

一方、コンピュータの処理速度の向上により、ディープラーニングが注目されている。その一つ的手法であるニューラルネットワークは脳神経系をモデルとした数理シミュレーションで、データの特徴を学習することによってパターンを認識し、未知のデータに対して分類を行う。我々はこのディープラーニングの技術を用いることで、注視点データから人間が画像

を認識する行動の特徴を特定できるのではないかと考えた。先行研究では、診断画像へのディープラーニングの適応は多く試みられているが、注視点を対象とした研究は少ない。

そこで我々は読影プロセスの評価を目的に、ディープラーニング技術を用いて、読影中の注視点軌跡における特徴量の分析を試みた。

2. 方法

2.1 対象

対象画像は10枚のCTコロノグラフィ腸管内部展開像とした。図1にCTコロノグラフィの臨床像と読影で指摘すべき特徴部位の一例を示す。読影用画像は事前に病変部を示されており、これらにはポリープや側方発育型腫瘍などの病変像に加え、バウヒン弁や残渣などの非病変像も描出されており、読影で確認すべき複数のポイントが含まれている。読影は8名の診療放射線技師で行った。それらの平均勤務経験年数は35±9年であった。なお、本研究は群馬県立県民健康科学大学における倫理委員会承認され、診療画像の使用許諾、参加者に対する倫理的配慮を行った上で実施された。

2.2 分析画像の準備

まず、読影者が1名ずつ、10枚の画像に対し、1枚あたり約30秒で順次読影を行い、注視点追跡データを記録した。ただし画面外の位置にある点は除外した。注視点追跡データの取得にはJVCケンウッド社製ゲイズファインダーを用いた。この装置では、注視点データが0.02秒毎にモニタのXY位置で記録することができる。

次に、連続した30点(0.6秒相当)の注視点を取出し、それらのデータの重心を中心とした224×224pixelの白色パッチ画像に対して、注視点の軌跡を描いた。この時、0.02秒の時間経過ごとに白から黒へと軌跡線の濃度を変化させ、軌跡の過程をグラデーションで表現した。この注視点軌跡画像を、1辺あたり1/8に圧縮し(28×28pixel)、かつ軌跡の線分に対

して3ピクセルの膨張処理を施した。図2に注視点軌跡画像の例を示す。この処理を読影者ごと、および読影画像ごとに、注視点点を1点ずつずらして、オーバーラップさせながら繰り返し取得した。これにより、79,939枚の注視点軌跡画像を取得できた。

最後に、トレーニングセットとテストセットを作成する。トレーニングセットはディープラーニングで学習のために使用され、テストセットは結果の検証のために用いられるものである。トレーニングデータセットとして、6枚のCTコログラフィ画像に基づく注視点軌跡画像を取り分けた。なお、学習結果に偏りが出ないようにするため、病変部を通る軌跡画像と病変のない部位を通る軌跡画像ごとに、それぞれ同数の画像をランダムに並べた。よってトレーニングセットは10,426枚で構成される。また、テストセットとして、残り4枚のCTコログラフィ画像に基づく注視点軌跡画像を取り分けた。なお、テストセットでは分析結果の過大評価を防ぐため、注視点軌跡がオーバーラップしないように考慮して画像を取得した。このテストセットは206枚で構成される。

2.3 ディープラーニングアルゴリズムと分析手順

ディープラーニングアルゴリズムとして、オートエンコーダを採用した。オートエンコーダはニューラルネットワークを用いた教師なし学習を提供するアルゴリズムである。オートエンコーダの階層構造は入力層、中間層、出力層の単純な3層構造とした。学習方法として、入力層に注視点軌跡画像を投入し、中間層でウェイトを計算させ、出力層で入力画像を再現させる。繰り返しウェイトを調整することによって再現性を高めるように学習が進められる。本研究ではプログラミング言語としてpythonを使用し、ディープラーニングライブラリとしてgoogle社が提供するtensorflowを用いた。本研究で用いたオートエンコーダのモデル図を示す(図3)。

分析手順として、まず、トレーニングセットを用いて、オートエンコーダの学習を行った。この時、中間層のニューロンの数を16とした。これは注視点の動きのパターンは「見た」か「見てない」だけとは限らず、未知の行動も分類される可能性があるため、広めに設定した。学習回数は10,000回とした。次に、トレーニング済みのウェイトを用いて、テストセットを分類した。この分類結果を基に、注視点の軌跡を原画像に重ねて表示し、どのような行動かを観察した。ここで複数人の注視点軌跡を総合的に評価するため、注視点軌跡に対してガウシアンフィルタを適応し、ノーマライズして表示させた。ただし、これは結果の解釈を促すためのもので、ディープラーニングにおいて、この処理は施していない。またウェイトごとに原画像で病変のある部位と病変のない部位を通る注視点軌跡画像を別々に集計して、どのような違いがあるかを推察した。

3. 結果と考察

16のウェイトを視覚化して図4に示す。視覚化されたウェイトでは、軌跡の通り道を思わせるようなライン状の様相が確認できた。このウェイトは、いわゆる画像処理におけるフィルタと同義であり、つまり入力画像からの特徴抽出フィルタとして動作する(以下、フィルタと呼ぶ)。

表1に、フィルタで分類されたそれぞれの画像数を示す。本報告ではこれより、件数が最大の6番フィルタに絞って議論を進める。図5に6番フィルタで分類された注視点軌跡を原画像に重ねて表示したものを示す。ここでは注視点エリアをはっきり見せるために、バックグラウンドの原画像を薄暗く、注視点エリアを明るく描画している。図5と図1は同部位であるが、これらを比較すると、注視点エリアの部位にはポリープやクリッ

プがあり、画像上の臨床的に意味のある特徴部位を認識している視線の動きであると推察できた。さらに病変がない部分でも注視点エリアの存在が観察できた。表1の病変部と非病変部に分けた集計でも、6番フィルタはどちらも最大であった。読影という行為を考えると、病変にかかわらず画像の特徴的な部分を良く観察することは必然であり、行動特性として数多く抽出されると考えられる。このことから、我々はこの6番フィルタで分類された注視点軌跡が画像を良く観察する行動であると判断した。

これらのフィルタを大別すると、注意深く観察している行動と画像を走査している行動に分けられる。このことは従来の注視点解析において、単に「見た・見ていない」という結果を、病変部に目を通したが気付かなかった「見落とし」や病変部を認識していながらも病変ではないと判断した「見逃し」という行動を別々に評価できる可能性がある。つまり、本手法によって読影プロセスを表せるのではないかと推察された。

本研究の限界として、まず対象画像数、参加者数ともに少数であることがあげられる。また得られた結果に対して、調査者が解釈を後付けしているため、人間では理解不能な微細な行動を表す結果に対して不適切な解釈を与えている可能性がある。さらに、読影者は診療放射線技師に限定しているため、放射線科医などでは異なる結果を示す可能性もある。これらは今後、サンプル数やバリエーションを増やし、検証する必要がある。

4. まとめ

本研究では読影プロセスを評価する目的で、診療放射線技師によるCTコログラフィの読影中の注視点軌跡を対象に、教師なしディープラーニング手法を用いて分類を行った。結果では、画像上の意味のある特徴部位を注意深く確認する注視点軌跡を抽出することができた。本研究の方法論を実装した評価システムが確立できれば、読影技術を適切に指導できる教育が実現可能である。

参考文献

- 1) 人工知能学会. 深層学習-Deep Learning-. 近代科学社, 2016.

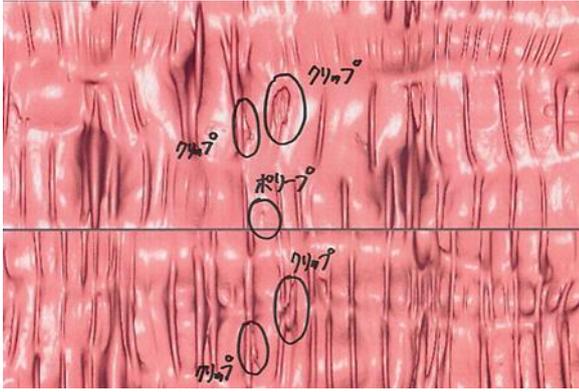


図1 CT コロノグラフィ

図は本研究で用いた CT コロノグラフィの臨床像と読影で指摘すべき特徴部位の一例である。

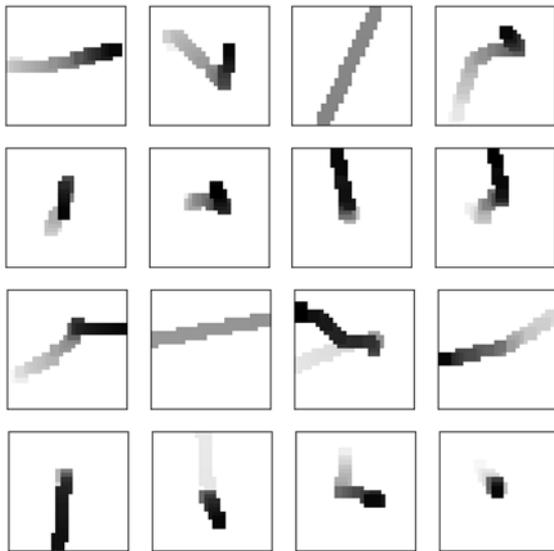


図2 注視点軌跡画像

図は本研究の注視点軌跡画像の抜粋である。28×28pixel の画像パッチに 3pixel の軌跡線が描かれ、時間経過ごとに白から黒へ濃度を変化させて描画している。

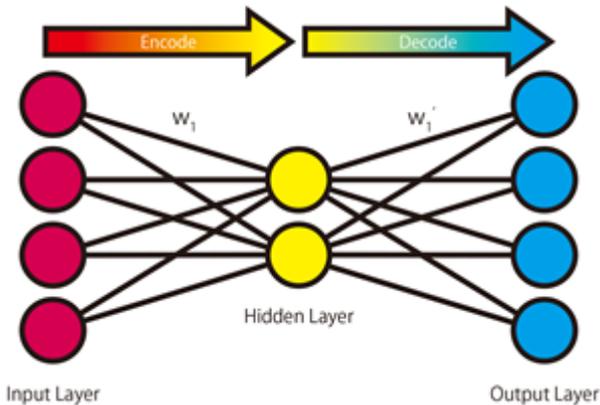


図3 オートエンコーダ

図は本研究で用いたオートエンコーダのモデル図である。入力層に画像が入力され、出力層で入力された画像が再現されるように中間層のウェイトが学習される。

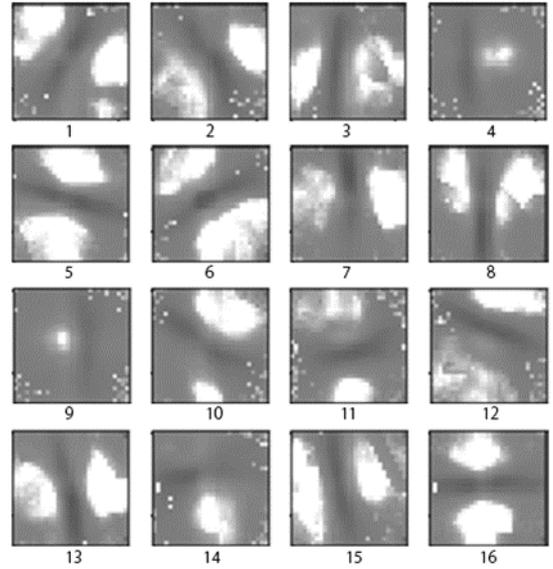


図4 ウェイトの可視化

図はオートエンコーダで学習されたウェイトを可視化したものである。

表1 フィルタごとのテストセットにおける分類数

フィルタ	病変部	非病変部	計
1	9	5	14
2	5	6	11
3	5	10	15
4	0	4	4
5	6	4	10
6	28	13	41
7	12	6	18
8	10	12	22
9	3	3	6
10	1	6	7
11	2	6	8
12	3	2	5
13	3	2	5
14	3	2	5
15	8	6	14
16	5	16	21
計	103	103	206

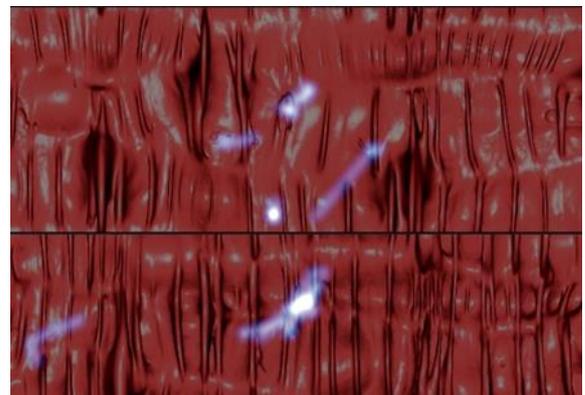


図5 注視点軌跡像

図は6番フィルタで分類された注視点軌跡を原画像に重ねて表示したものである。注視点軌跡は複数人のデータがあるため、軌跡の集中に基づいてガウス分布で表示している。