

機械学習による生体画像解析について

田中敏幸

慶應義塾大学 理工学研究科

1. はじめに

近年、機械学習を用いた画像解析手法の医療や美容の分野への利用が注目されている。システム開発の容易さなど機械学習の利点が強調された研究が多いが、機械学習の欠点のためにある程度以上の成果が出ていないのが現状である。本稿では、肌画像解析[1]-[5]や病理画像解析[6]-[8]を例として、機械学習のどのような点が問題となっているのかについて説明を行う。

2. 肌画像解析

2.1 利用する肌画像

解析に用いる画像は、Fig. 1に示すスカラー社製マイクロスコープで撮影を行う。画像は1280×960画素のJPEGフォーマットのRGB画像として保存される。また、対象者は20歳から80歳の成人女性とする。



Fig. 1 入力デバイス（スカラー社製）

Table 1 画像と対象者

使用機器	Scalar社のUSBマイクロスコープ+25倍拡大レンズ
撮影サイズ	1280×960画素
対象者	成人女性(M=20)
年齢	20-80歳

2.2 解析のためのAIシステム

肌画像解析を行うニーズとしては、肌の性質に合った化粧品の選定、肌アレルギーに対するアレルギー物質の選定などがあげられる。機械学習を用いてこのようなシステムを構築する場合、Fig. 2のようなものが一般的に考えられている。

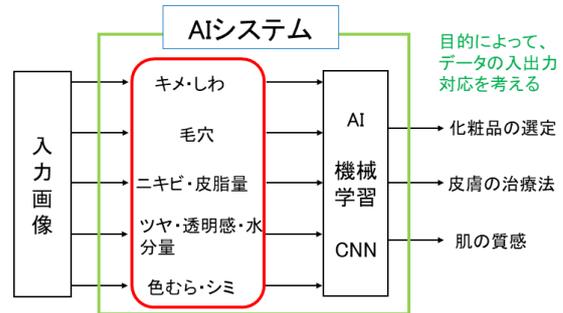


Fig. 2 肌画像解析のための標準的AIシステム

最近では、畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) あるいは深層学習 (Deep Learning) と呼ばれるものが利用されている。しかし、深層学習を利用する場合には、次のような点が問題となる。

- (1) 非常に多くのデータセットが必要
 - (2) 確からしさが十分なデータセットが必要
 - (3) 入出力の対応に正解がなければならない
- データの数を集めるだけであれば時間をかければよいが、データの質や正確な対応関係については、十分なものが得られない可能性がある。肌画像解析の応用では、Fig. 3のように特徴量解析と機械学習を分けることによって、データセット数が多くなくても対応できるようにできる。

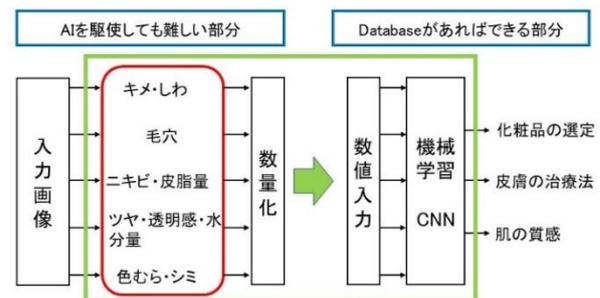
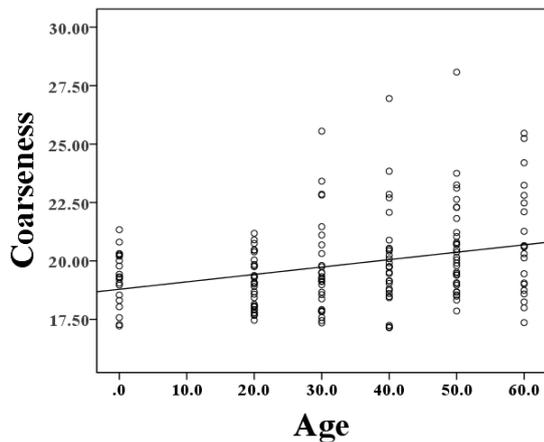


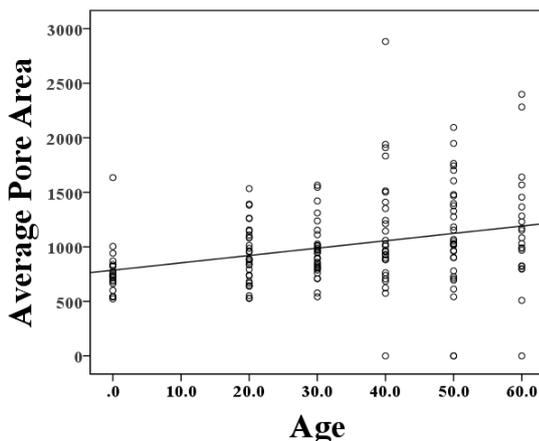
Fig. 3 画像解析を分離した肌解析AIシステム

2.3 肌画像解析の問題点

しかし、確からしさが十分な正確な対応関係という点ではどうだろうか。一例として、肌のキメの細かさと毛穴の大きさを計測した結果をFig. 4に示す。キメや毛穴は肌の質感に大きく影響するパラメータとなっている。



(a) 肌のキメの細かさ



(b) 毛穴の大きさ

Fig. 4 肌画像解析から得られる指標

Fig. 4は各年代における計測結果を表しているが、非常にばらつきが大きいこと、相関が非常に小さいことがわかる。肌の質感は年齢とともに変わると言われているが、年齢で分けたデータセットでは、その後の機械学習がうまくいかないことが容易に想像できる。

このように、人を対象とした機械学習は、学習セットそのものの作り方が難しく、学習セットによって行うことができるサービスが限定されるという問題もある。

3. 病理画像解析

3.1 病理画像の背景と目的

次に病理画像解析について、機械学習の利用方法を見ていくことにする。病理医の中では、遺伝子変異が癌細胞の組織形態に影響を与えているという仮説がある。実際、改訂ベゼスダガイドラインという治療ガイドラインにおいて、ミスマッチ修復遺伝子変異が原因とされるリンチ症候群では、腫瘍内リンパ球浸潤、髄様増殖、粘液癌・印環細胞癌様分化、クローン様リンパ球反応などの組織学的特徴があるとされている。この仮説が正しいとすれば、病理画像の画像特徴から遺伝子変異の有無を予想することができる。これが可能となれば、先述した現状の遺伝子検査のデメリットを改善し、検査期間の短縮、医療費の削減、病理検体の限局につながる。

3.2 利用する画像とCNN

ここでは、国立がん研究センター東病院より提供された検体を用いて解析を行った。これらは、HE染色された生検材料及び手術材料の顕微鏡画像を、バーチャルスライドスキャナによってデジタル画像化したものであり、階調数は256、色空間はRGB色空間のJPEG画像である。オリジナルのバーチャルスライドは、20倍×40倍の解像度が226nm/pixel (112389dpi)である。本研究ではデータ容量を削減するため、20倍×20倍とした画像を用いた。遺伝子検査の結果に関しても検体の画像との対応も含めて同病院から提供を受けている (Table 1参照)。これらの検体と機械学習によってMSI検査による遺伝子変異陽性と陰性の分類を行った。

Table 1 MSI検査結果に対応する材料数

	MSI陽性	MSI陰性
手術材料	40	40
生検材料	20	20

研究では、学習済みCNNとして、2014年のISLVRCにおいて優秀な成績を残したVGG19モデルを用いて転移学習を行った。このモデルは16層の畳み込み層と5つのプーリング層、3つの全結合層で構成されており、本研究では、入力をHE染色画像、出力を遺伝子検査結果予測の2値分類として、出力側の全結合層を作り変えた(Fig. 5)。

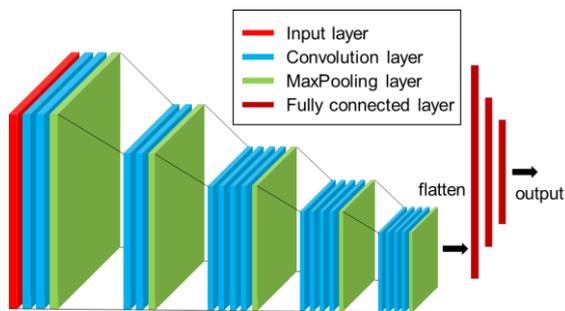


Fig. 5 利用するCNN (VGG19モデル)

3.3 病理画像解析の問題点

VGG19によってMSI検査結果を予測した結果をTable 2に示す。3回の試行に対して、CNNの性能評価によく用いられるAccuracy, Precision, Specificity, F-measureを示している。Specificityの結果はよいとは言えないが、他は80%前後の結果となっている。

Table 2 HE染色画像からのMSI検査予測結果

	Accuracy	Precision	Specificity	F-measure
1回目	0.796	0.812	0.649	0.824
2回目	0.730	0.712	0.500	0.794
3回目	0.734	0.760	0.477	0.811
平均	0.744	0.761	0.542	0.810

CNNの応用としては興味深い内容で、簡単な実装である程度の結果が得られている。しかし、この結果では実際の臨床現場で利用することはできない。十分な結果が得られていない、原因として次のようなことが考えられる。

- (1) 細胞画像に見られる特徴は個人差が大きい
 - (2) HE染色画像とMSI検査結果をどのように対応させるかが難しい (画像全体を対応させて良いのか)
 - (3) CNNは何を見て判定しているのか不明
 - (4) 医師が判定できない資料を利用している
- 特に、上記の(3)については、特徴量マップを調べることによってHE染色画像のどこをとらえたのかを確認した。医師に評価してもらったところ、特徴量マップとMSI検査結果との対応理由が不明とのことであった。

4. 機械学習に用いる生体画像

美容や医療への応用のために生体画像を利用する場合、画像や対応データにいろいろな制約があるので、そのことについてまとめておく。

- (1) データのばらつきを抑える

生体画像で個体ごとのばらつきがあるのはやむを得ないが、ばらつきが少なくなるようにグループを分けるなどの工夫が必要となる。肌の解析の際に説明したように、年齢によるグループ分けが正しいとは限らない。

(2) 正確な対応関係

病理画像の際に説明したように、入力と出力の対応がはっきりしていることが重要である。医師との共同研究においては、医師は最終的に新しい知見を得たいと考えている場合が多い。そのような場合には、正確な対応関係のわからないデータを扱う場合もある。どのような対応関係としてまとめるかも工夫が必要となる。

生体計測にCNNを用いる場合、ある程度の成果は短時間で得られるが、そこから先については非常に難しい問題となる。

5. おわりに

本稿では、美容関係と医療関係の研究においてCNNや深層学習を用いた時、何が難しいのか、なぜ難しくなるのかを例を用いて示した。機械学習は多くの分野で期待されている技術であるが、必ずしも万能ではない。どのような点に気を付けるかだけでなく、機械学習を行いやすい形に研究そのものを変更したり、サービスを変更したりすることも含めて考える必要がある。最初から最終目的が決まっている場合、機械学習を選ぶかどうかも含めて検討が必要である。

参考文献

- [1] Wu Y., Akimoto M., Igarashi H., Shibagaki Y., Tanaka T. : Quantitative Assessment of Age-dependent Changes in Porphyrins from Fluorescence Images of Ultraviolet Photography by Image Processing, Photodiagnosis and Photodynamic Therapy, Vol. 35, (2021)
- [2] Wu Y., Tanaka T. : Objective and quantitative measurement of skin micro-relief by image analysis and application in age-dependent changes, Skin Research and Technology, (2021)
- [3] Yue Wu, Toshiyuki Tanaka, Makio Akimoto : Utilization of Individual Typology Angle (ITA) and Hue Angle in the Measurement of Skin Color on Images, Bioimages, Vol. 28, pp. 1-8, (2020)

- [4] Wu Y., Yi F., Akimoto M., Tanaka T., Meng H., Dong Y. : Objective measurement and comparison of human facial skin color in East Asian females, *Skin Research and Technology*, Vol. 26, No.4, pp. 584-590, (2020)
- [5] 田中敏幸：画像解析による特徴量抽出に基づいた AI 肌診断システム、第 12 回化粧品開発展（アカデミックフォーラム）、(2022)
- [6] 藤原菜穂、田中敏幸：畳み込みニューラルネットワークによる胃癌の悪性度判別、パーソナルコンピュータ利用技術学会論文誌、Vol. 13, No. 1, pp. 31-36, (2019)
- [7] 藤井琢也、田中敏幸：画像処理と深層学習によるマイクロサテライト不安定性の予測、第16回パーソナルコンピュータ利用技術学会全国大会、A2-3、(2021)
- [8] 佐藤善彬、田中敏幸：深層学習と色差解消手法による肺癌生検組織標本の組織型分類、第 16 回パーソナルコンピュータ利用技術学会全国大会、A2-4、(2021)