

皮膚癌画像診断のための生理的知見に基づく前処理可用性の判定

岡本秀明 彌富仁

法政大学 理工学部 応用情報工学科

概要

メラノーマは予後不良の皮膚癌で早期発見が極めて重要だが診断は困難なため、様々な自動診断システムが提案されてきた。我々は、母斑（良性）とメラノーマの発生過程の医学的知見に基づき、腫瘍領域の長軸位置合わせを行うことで、少ない学習データからでも頑健で良好な診断が可能であることを確認した。しかし、境界不明瞭な腫瘍や影などの多い悪条件症例には上記手法が使えず、より大規模で複雑な解析を要する。そこで本報告では皮膚腫瘍画像の長軸位置合わせの可用性を深層学習器で推定し、後段での識別を振り分ける手法を提案する。メラノーマを含む7種の皮膚腫瘍の識別問題において、提案手法は従来手法と比較して識別精度が3.8%向上した。

1. 背景

皮膚癌の一種であるメラノーマは、特に進行が早く早期発見が極めて重要である。診断はダーモスコピー（偏光レンズがついた皮膚科専用の拡大鏡）を用いて行われるが、熟練の専門医による診断においても診断精度は75-84%だと言われているほど難しい[1]。加えて診断は専門医の主観に依存するため、再現性も乏しい。このような背景から、メラノーマの早期発見の支援および客観性のある診断の実現を目的とし、画像認識の技術を用いたメラノーマ自動診断システムの研究が進められてきた[2]。畳み込みニューラルネットワーク(CNN)における診断システムでは極めて多くの学習データを必要とする欠点があるが、従来困難であった前処理などを不要としながらも、高精度な識別が報告されている[3]。一方で、医学的知見に基づく母斑/メラノーマの形状に着目した前処理(長軸位置合わせ)を行うことで、それぞれが持つ形状の差異を顕著化し、少ない学習データからでも良好な診断精度の実現が報告されている[4]。しかし境界不明瞭な腫瘍や、体毛や泡などのノイズ、影などが多く腫瘍領域の獲得が困難な悪条件症例の場合、位置合わせは困難を伴う。またこうした症例の適切な診断には大規模で複雑な解析を要する。

本報告では、皮膚腫瘍画像の長軸位置合わせの可用性を深層学習器で推定し、その結果に応じた後段の各病例識別

器で識別を行う手法を提案する。好条件症例は小規模で十分高い識別精度を持つ、位置合わせを行った画像のみを学習した識別器を使用し、悪条件症例は位置合わせを行っていないより大規模な画像を学習した識別器を使用する。これにより全体の診断精度の向上を目指す。

2. 方法

2.1 長軸位置合わせ

母斑の細胞は一定の秩序を保ちつつゆっくりと成長するため形状が左右対称となる場合が多いのに対し、メラノーマの細胞は無作為かつ盛んに成長するため形状が左右非対称になる場合が多く見受けられる。吉田らは、このような形状の差異に着目し、前処理として腫瘍領域の長軸位置合わせを提案している。腫瘍領域の長軸位置合わせを行うためのアルゴリズムは、(1)腫瘍領域の決定、(2)長軸の決定、(3)回転/リサイズから構成される。長軸位置合わせのイメージを図1に示す。

2.2 位置合わせ分類器

位置合わせの可用性を深層学習器で推定する前処理を導入し、その結果に応じて2つの病例識別器の使い分けを行う。位置合わせ分類器使用時のシステム構造を図2に示す。位置合わせ分類器の分類結果により入力画像が位置合わせ可能ならば病例識別器Aに送り、位置合わせ不可能ならば病例識別器Bに送る。2つの識別器は同じCNNモデルを使用するが、表1に示す通り学習データと位置合わせの有無が異なる。

2.3 データセット

本報告では、表2に示す7種類の皮膚腫瘍から構成されるISIC 2018 task 3データセット[5]、計10,015枚のダーモスコピー画像を使用した。これらのうち、上記の長軸位置合わせ処理が適切に行えたものは6,450例であった。

2.4 システムの構成

位置合わせ分類器はImageNetで事前学習済みのCNNであるVGG16モデルをベースとし、長軸位置合わせ処理の可否を教師データとして全結合層のみを再学習することで構築した。出力層はシグモイド関数を有する1ノードとし、最も分類精度が高くなるような閾値を選択した。学習パラメータは学習率を0.001、学習回数を20回、バッチサイズを20とした。

病例識別器A、Bは、それぞれ表1に示した条件のもと構築した。これらも同様にImageNetで事前学習済みのVGG16モデルをベースとし、全結合層のみを再学習した。

Determination of preprocessing availability based on physiological finding for image-based skin cancer diagnosis
Hideaki Okamoto, Hitoshi Iyatomi (Applied Informatics, Faculty of Science and Engineering, Hosei University)

学習パラメータは学習率を 0.001, 学習回数を 50 回, バッチサイズを 20 とした. いずれの識別器も 5-fold cross validation にて評価した.

3. 実験と結果

位置合わせ分類器において, 全データ 10,015 例のうち分類を誤ったデータは合計 481 例であり, 位置合わせ分類器単体の分類精度は 95.2%であった. 誤りの内訳は, 可能を不可能と誤ったデータは 464 例, 不可能を可能と誤ったデータは 17 例であった.

続いて位置合わせ分類器の有無による病例識別精度の違いの検証を目的とした比較実験を行った. 全ての学習データから構成される病例識別器 B での評価をベースラインとする従来手法と, 提案手法である位置合わせ分類器の結果に応じて病例識別器 A, B を使い分ける手法において 7 クラス balanced accuracy (各クラスの識別率の平均値: 以下識別率) を比較した.

その結果, 従来手法における識別率は 81.7%であり, 提案手法では同 85.5%と 3.8%の向上が見られ, 位置合わせ分類を事前に行うことの有効性が確認できた.

提案手法において病例識別器 A に送られたデータは 6,003 例, 識別率は 87.3%であり, 識別を誤ったデータは 762 例, 識別・位置合わせ分類共に誤ったデータは 15 例であった. 病例識別器 B に送られたデータは 4,012 例, 識別率は 82.8%であり, 識別を誤ったデータは 690 例, 識別・位置合わせ分類共に誤ったデータは 123 例であった.

4. 考察と今後の課題

提案手法による識別率は, 全てのデータを学習した病例識別器 B のみを用いるベースラインよりも良好な識別能を示した. これは病例識別器 A が長軸位置合わせによる恩恵を直接受けたこと, 病例識別器 B が位置合わせ不可能データだけに集中することで何らかの恩恵を受けたことが要因



図 1: 長軸位置合わせ

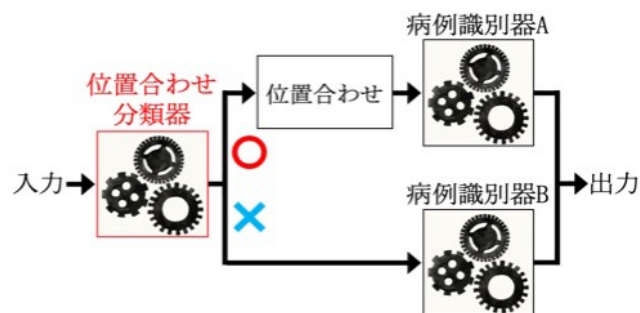


図 2: 位置合わせ分類器使用時におけるシステム構造

表 1: 病例識別器 A, B の相違点

	病例識別器 A	病例識別器 B
学習データ [†]	位置合わせ可能のみ (6,450 例)	全データ (10,015 例)
学習・評価時の位置合わせ	あり	なし

[†] 5-fold cross validation のため, 各 fold では全体の 4/5

表 2: 実験で用いたデータセットの詳細

MEL	NV	BCC	AKIEC	BKL	DF	VASC	Total
1,113	6,705	514	327	1,099	115	142	10,015
位置合わせ可能			位置合わせ不可能				
6,450			3,565				

であると考えられる.

誤って病例識別器 A に送られた位置合わせ不可能データは 17 例あり, そのうち 15 例の病例識別を誤った. 識別率が極端に低くなった要因は, 位置合わせ不可能データに位置合わせ処理を行うことで画像上の腫瘍領域が大きく崩れたためと考えられる. 病例識別器 A における位置合わせ分類を誤った(不可能→可能)データの識別率が低い一方で, 逆の誤り(可能→不可能)は病例識別器 B での精度にほぼ影響しないため, 閾値は位置合わせが確実にできる値に設定する必要がある. また位置合わせ分類器自体の精度向上も今後の課題として挙げられる. 深層学習に頼らない特徴量抽出と深層学習による特徴量抽出の併用等を検討することで分類精度向上を目指す.

参考文献

- [1] G. Argenziano, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera et al., "Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet," *Academy of Dermatology*, vol. 48, no. 5, pp. 679-693, 2003.
- [2] H. Iyatomi, H. Oka, M. E. Celebi, M. Hashimoto, M. Hagiwara, M. Tanaka et al., "An Improved Internet-based Melanoma Screening System with Dermatologist-like Tumor Area Extraction Algorithm," *Computerized Medical Imaging and Graphics*, vol. 32, no. 7, pp. 566-579, 2008.
- [3] A. Esteva, B. Kuprel, R. A. Novoa et al: Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, vol.542, pp.115-118, 2017.
- [4] T. Yoshida, M. E. Celebi, G. Schafer, H. Iyatomi, "Simple and effective pre-processing for automated melanoma discrimination based on cytological findings," *IEEE Proc. Big Data*, pp. 3439-3442, 2016.
- [5] P. Tschandl, C. Rosendahl, H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions", *CoRR*, abs/1803.10417, 2018.