法政大学学術機関リポジトリ

HOSEI UNIVERSITY REPOSITORY

PDF issue: 2024-12-22

転移学習を用いたメラノーマ自動診断システ ム

吉田, 琢也 / Yoshida, Takuya

(出版者 / Publisher) 法政大学大学院理工学・工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要.理工学・工学研究科編 / 法政大学大学院紀要.理工学・工 学研究科編

(巻 / Volume)
58
(開始ページ / Start Page)
1
(終了ページ / End Page)
6
(発行年 / Year)
2017-03-31
(URL)
https://doi.org/10.15002/00014237

転移学習を用いたメラノーマ自動診断システム

AUTOMATED MELANOMA DIAGNOSIS SYSTEM WITH TRANSFER LEARNING USING DEEP CONVOLUTIONAL NEURAL NETWORK

吉田 拓也
 Takuya YOSHIDA
 指導教員 彌富 仁

法政大学大学院理工学研究科応用情報工学専攻修士課程

In recent years, a new machine learning schema called deep learning has gained many promising achievements in a wide range of industries and thus attracted more than just researchers. A convolutional neural network (CNN) is a principal aspect of deep learning techniques specialized for machine learning and/or computer vision. Through the influence of these promising achievements, replacement of existing methods has been progressed rapidly. The automated melanoma diagnosis system is no exception, and an automated melanoma diagnosis system using CNN has already been proposed. However, CNN's learning requires a large amount of training data, and since melanoma images can only be obtained from limited organizations such as medical institutions, it is difficult to collect a large amount of data. In this study, we overcome these problems using transfer learning and applying effective pre-processing to the data set. The melanoma classifiers constructed using the proposed method achieved AUC: 0.873 on the ROC curve. *Key Words : melanoma, skin lesion, dermoscopy, convolutional neural network, transfer learning*

1. はじめに

メラノーマは悪性度の高い皮膚癌であり、特に初期は 外観が母斑に酷似していることから診断が難しい.メラ ノーマの5年生存率は0期や1期であれば90%以上,II期 でも70-80%であるが、III期では50%弱、IV期では10% 未満と予後は格段に悪くなる[1].

そのため、初期のメラノーマを判別するために、表皮 構造を詳細に観察する非侵襲的な表皮造影技術であるダ ーモスコピー [2]が導入された.また同時に、ダーモスコ ピーの診断に基づく診断指標が開発された(ABCD-rule [3],7-point checklist [4]等).しかし、ダーモスコピーに よる診断はしばしば主観的なものであるため、正確な診 断には豊富な経験とトレーニングが必要であり、熟練し た皮膚科専門医でも正答率は75-88%にとどまる [5].こ のような背景から、視覚的解釈の困難さと主観性に起因 する診断誤差を最小限にするために、メラノーマの自動 診断システムの開発が進められてきた [6-8].

従来の自動診断システムは、I) 腫瘍領域抽出、II) 特 徴量設計・抽出、III) 機械学習による分類 の3つのフェ ーズから構成されるが、診断に用いられる症例には、腫 瘍領域に体毛がオーバーラップしているものや、光の乱 反射を防ぐために用いられる超音波ジェルにより気泡が 生じているものなど、腫瘍領域抽出を行う際に障害とな るものが映り込んでいる症例が多く存在する.このよう な処理が難しい症例の様々なパターンに対応できる汎用 的な腫瘍領域抽出手法,および特徴量設計は未だ確立し ておらず,ゆえに自動診断システムでは扱える症例は限 られていた.

一方,近年では音声認識や画像認識のベンチマークテ ストで、高い識別精度を実現したことから、深層学習 (Deep Learning)の有効性が広く認知されるようになった [9,10]. 深層学習とは、従来のニューラルネットワークよ りも多層のニューラルネットワークを意味するが、一口 に深層学習と言っても、様々なアプローチがある.画像 認識分野において主流でありベースとなるのが Convolutional Neural Network (CNN) [11]である. CNN に よる特徴量抽出は学習の一環として入力画像の生の画素 値から直接行うため、従来システムのボトルネックとな っていた,「腫瘍領域抽出」,「特徴量設計」「特徴量 抽出」などの複雑な処理を伴わずに分類を行うことがで きる. そのため, 近年では CNN を用いたメラノーマ自動 診断システムが提案され始めている [12-14]. CNN では 上記の理由から扱える症例には制限がないという利点が あるが、しかし一方で、CNN では有効な特徴を学習する ために,膨大な学習データを要するという問題があった. そこで、本研究では、ドメインの異なる大規模データ

セットで事前に学習済みのモデルを用いた Fine-tuning に よる転移学習 [15]を行うことで学習効率の向上を図った. また,学習で用いるデータセットには,一般的な data augmentation に加え,医学的知見に基づいた前処理を施す ことで,少規模データセットでも十分に効果的な学習が 行えるようにした.

2. 体毛除去

我々の提案する前処理は、後述する腫瘍領域特定フェ ーズにて色調情報を用いた簡単な閾値処理を伴う.この 処理において、腫瘍領域に体毛がオーバーラップしてい る症例においては腫瘍領域の特定が難しい.そのため、 長軸の位置合わせ処理を行う前に体毛を適切に除去する 必要がある.体毛の除去は(1)体毛抽出,(2)画像修 復の2つのフェーズに分けられる.

(1)体毛抽出:

体毛抽出の手法として本研究では、Frangi および Sato らが提案したヘッセ行列に基づく線状構造抽出を行う multi-scale フィルタ [16,17]を用いる. multi-scale フィルタ は、入力画像に二次元ガウシアンフィルタを適用し、雑 音除去を行うと同時に二次微分によって体毛のような線 状構造を強調するフィルタである.線状構造では、線の 断面方向への濃淡変化が大きく、線の進行方向への濃淡 変化は小さい.したがって、線状構造を強調する場合の ヘッセ行列の固有値 λ_1, λ_2 ($\lambda_1 > \lambda_2$)の関係式は式(1)のよ うになる.

$$\lambda_2 << \lambda_1 \approx 0 \tag{1}$$

これは、ヘッセ行列の固有値λ₁, λ₂の値がこの関係式に近いほど、画素の線状構造が強いことを表す. 各画素の線 状性を式(2)で定義する.

$$\lambda_{12} = \begin{cases} |\lambda_2| (1 + \lambda_1 / |\lambda_2|) = |\lambda_2| + |\lambda_1| & (\lambda_2 \le \lambda_1 \le 0) \\ |\lambda_2| (1 - \alpha \lambda_1 / |\lambda_2|) = |\lambda_2| - \alpha \lambda_1 & (\lambda_2 < \lambda_1 < |\lambda_2| / \alpha) \\ 0 & (otherwise) \end{cases}$$
(2)

ここで、 α の値は、 $0 \le \alpha \le 1$ で、線の双極性に対する許 容度を表す.この λ_{12} の値が大きいほど、線状構造に近い ことを表す. Fig. 1 (b)に体毛抽出の結果を示す.

(2)画像修復:

前フェーズで抽出した体毛をマスクとして用いて、マ スクにより選択された体毛領域の除去および除去部の修 復を行う. 画像修復の手法は,Bertalmio らが提案した Image Inpainting [18]のアルゴリズムを用いた. Image Inpainting は濃淡値の連続性を用いた手法であり、欠損領 域の周りから濃淡値を滑らかに補間することで欠損領域 を修復する. 画像修復の結果を Fig. 1 (c)に示す.



Fig.1 腫瘍領域の体毛除去

3. 長軸の位置合わせ処理

長軸の位置合わせ処理のフローを Fig. 2 に示す. この処 理は(1)初期領域特定,(2)領域選択,(3)長軸の計 算・位置合わせの3つのフェーズから構成される.



Fig.2 長軸の位置合わせ処理フロー

(1) 初期領域特定

このフェーズでの処理は Iyatomi ら [7]の提案手法に基 づき,まずガウシアンフィルタを用いた雑音除去を行う. その後,ラプラシアンフィルタを用いて濃淡変化の大き い画素値の座標集合を取得した.集合に含まれる画素に 対する B(青)の成分を用いて,大津の閾値判別法によ り閾値を決定し,暗い部分を初期腫瘍領域とした.

(2)領域選択

前フェーズの閾値処理により,無数の独立した小領域 が生成されるため,このフェーズでは最初にそれぞれを ある程度の大きさを持つ領域にまとめる処理を行う.す べての小領域に対してラベリング処理を行い,各領域に おいて,画像全体の1%よりも小さい領域は,自領域より も大きく最も長い境界線を共有する領域に統合する.こ の処理をすべての領域に対して領域数が十分小さくなる まで繰り返し行った上で,なお残った腫瘍領域候補に対 して,次は腫瘍領域とそれ以外の領域とを区別する処理 を行う.この処理は、ラベリングされた腫瘍領域候補の 中から,端に触れていない最も大きな領域を腫瘍領域と みなし,注目領域と比較し,注目領域が腫瘍領域の20% 以上の大きさを有すれば,前フェーズで濃淡値が低く閾 値処理で抽出できなかった腫瘍領域とみなして統合する.

(3) 長軸の計算・位置合わせ

このフェーズでは、まず腫瘍領域の重心を算出する. 重心は幾何学的に、ある図形の、その周りでの一次モー メントが 0 であるような点を重心と定義した.本研究で は、腫瘍領域の輪郭上の 2 点を結ぶ線分が、この重心付 近を通過し、かつ最長となるとき、その線分を腫瘍領域 の長軸として定義した.長軸が決定したら、この長軸が 水平方向に沿うように回転させ位置合わせを行う.後述 する CNN では正方行列の画像を入力に取るため、位置合 わせ後は腫瘍領域を中心に正方行列となるようトリミン グを行う.以上の長軸位置合わせ処理は、学習用、評価 用問わず、すべての入力画像に対して行われる.

4. Convolution Neural Network

Convolutional Neural Network (CNN) は, 複数の Convolution 層, Pooling 層, LCN (Local Contrast Normalization) 層と全結合層から構成される順伝播型ネ ットワークである.学習は誤差逆伝播法(Back-Propagation) と確率的勾配降下法 (Stochastic Gradient Descent) により 最適化を行う.

Convolution 層では、入力画像に対し、複数個のフィル タを並行して畳込み演算を行う. 畳込み演算の結果得ら れる中間層での出力は特徴マップと呼ばれる. Fig. 3 に Convolution 層での処理の概要を示す.



Fig. 3 Convolution 層

Convolution 層の出力を入力として扱う Pooling 層では, 入力された特徴マップ上に局所領域をとり,領域内の値 を統計し集約する処理を行う.正方領域内の画素値の決 め方はいろいろあるが,今回は領域内の最大値を出力と する max pooling を用いた. max pooling の処理の概要を, 以下の Fig.4 に示す.このように入力情報の一部を切り捨 てることで,各層間の結合(重み)に過疎生を持たせる ことができ,入力画像の位置ずれなどの幾何学的変化に 対して不変性が得られる.この処理により上層へいくほ ど広い受容野の情報が抽象化されるため,ゆえに不変性 は層を経るごとに強くなる.



Convolution 層と Pooling 層を経た後は,LCN 層で全チャネルにわたる局所領域ごとに画素の平均を求めて減算, そこから標準偏差を求め,除算する.この操作により, 画像上の局所領域内の各画素を平均0,分散1になるよう に濃淡が正規化される.

全結合層はここまでの特徴抽出部から最終的に出力さ れた特徴マップを入力にとり,各層間の重みは全結合さ れる.前層からの入力は一次元ベクトルとしてみなされ, 全結合層の各ユニットは線形和h_jで表され,最終出力は 式(3)に示すソフトマックス関数を用いて算出した全結合 層のユニットに対してクラス分類を行う.y_iが最大値をと るユニットを推定クラスとする.

$$y_i = \frac{\exp(h_i)}{\sum_{i}^{n} \exp(h_i)}$$
(3)

5. 転移学習

転移学習(Transfer learning)は、あるドメイン(元ド メイン)の知識を解きたい問題のドメイン(目標ドメイ ン)の学習に利用する手法である.学習データが非常に 少ない、あるいは全く存在しない場合でも、元ドメイン の知識を利用して分類器を学習できるという特性がある.

一般に、学習済みモデルを目標ドメインの特徴抽出器 として用いるアプローチと、目標ドメインのデータセッ トを用いて再学習を行うアプローチがある.後者は特に Fine-tuning と呼ばれ、学習済みモデルの識別層、あるい はその1つ前の Convolution 層だけを目標ドメインに置き 換えて、誤差逆伝播法による再学習を行う手法である. その他の部分は学習済みモデルのパラメータをそのまま 初期値として用いることが多い. CNN の学習は初期値依 存性が強く、特に学習データが少ない場合はできるだけ よい初期値を得ることが過学習を防ぎよい学習結果を得 るために重要である.そのため、学習済みモデルを適切 に選択し、初期値として用いた転移学習を行うことで、 フルスクラッチから学習するよりも効率的に学習を進め られる場合が多い.

6. 実験

本実験は、学習済みモデルのパラメータを初期値とし て用いて、そこから Fine-tuning を行い、メラノーマ/母 斑の2クラス分類器を構築した.また、比較実験として、 小規模な CNN を用いた同じデータセットによるフルス クラッチでの学習を行った.各実験の詳細を以下に述べ る.

(1)データセット

本研究では、2 つの医療機関(University Federico II of Naples, Italy; University of Graz, Austria) およびオープンデ ータセット (International Skin Imaging Collaboration (ISBI) [19])から入手したダーモスコピー像で構成される 1,760 枚(メラノーマ:329枚、母斑:1,431枚)を学習用デー タとして用いた.ここで、3つの機関から入手した病変の ラベルは、すべて生検の病理組織学的検査に基づいて診 断が確定されたものである.これらすべての症例は前処 理として,前述した長軸の位置合わせ処理に加え,一般 的な data augmentation を行った. 実施した data augmentation は順に、I) 垂直方向・水平方向への反転処 理, II) ガウス雑音の付加の2つである. ガウス雑音は 各画素の輝度値を中心に標準偏差0-50の範囲でランダ ムに 5 段階の雑音を付加する. 以上の data augmentation により,元のデータセットを20倍の35,200枚まで拡張し た. また, 各機関の画像は, 画像サイズにばらつきがあ るが、長軸の位置合わせ処理の段階で腫瘍領域を中心に 正方行列にトリミング,および 256×256 pixel にリサイズ した.

(2) Fine-tuning による学習済みモデルの再学習

本実験では、学習済みモデルに、ILSVRC 2014 で 2 位 の成績を収めた VGG-19 [9]を用いた. VGG-19 は、重み のある層(Convolution 層, 全結合層)を19層持つ深いネ ットワークであり、3×3の小さなフィルタを用いた畳込 みを連続して行っているのが特徴である.実験で使用し た VGG-19 は ImageNet [20]を用いて学習されている. ImageNet の 1,000 クラスには、メラノーマおよび母斑の クラスは含まれていないため, 元ドメインと目標ドメイ ンのカバーする領域は異なることが想定される. そのた め、今回の実験では VGG-19 を単に特徴抽出器として用 いるのではなく、Fig.5に示すように、VGG-19の全結合 層手前の Convolution 層以降のパラメータを Fine-tuning す ることで分類器の構築を行った.また, CNN では浅い層 ほどエッジやブロブなど汎用的な特徴が抽出されるのに 対し、深い層ほど学習データに特化した特徴が抽出され る傾向を考慮して、最終 Convolution 層より前の浅い層の パラメータに関しては、初期パラメータのまま固定し、

深い層の重みのみを,メラノーマ/母斑の二値分類に合 うように再調整した.

本実験は、Caffe フレームワーク [21]を用いて行った. 学習済みのモデルを再学習するため、全体の学習率は 0.001 と低く設定した.ただし、Fine-tunig を行う最終 Convolution 層と全結合層は新規にメラノーマ/母斑の 2 クラスを学習するため、重みとバイアスの学習率の乗数 は高い値に設定した.



Fig. 5 VGG-19 を用いた Fine-tuning

(3) CNN によるフルスクラッチ学習

比較実験のため,事前学習なしの CNN においても同じ データセットを用いて実験を行った.本実験ではメラノ ーマ 329 枚, 母斑 1,431 枚と比較的小規模なデータを扱う ため, CNN の構成は, Convolution 層と Pooling 層が 3 層 連なった小規模なモデル構成とした.具体的な構成とし て、まず第1層目の畳込み層は 256×256 の入力画像がラ ンダムクロップされた 224×224×3 を入力として取り, サ イズ11×11×3の96個のフィルタを用いて畳込みを行う. 第2層目の畳込み層は, 第1層目のレイヤの出力を入力 として取り、サイズ7×7×96の48個のフィルタを用い て畳込む. そして第3層では, 第2層目のレイヤの出力 を入力として取り、サイズ5×5×48の128個のフィルタ を用いて畳込む. 第3層目のレイヤの出力は全結合層へ と渡される. 各レイヤの畳込みの出力に対しては、サイ ズ3×3のフィルタを用いた max pooling, および局所コン トラスト正規化を行った.

7. 結果

学習は、バッチサイズ 100 のミニバッチ学習に基づく 確率的勾配降下法で行い、損失関数に交差エントロピー 誤差関数を用いた.学習回数は 100 epoch であり、各ミニ バッチ処理の終了時に重み、およびバイアスが更新され る.両分類器共、100 epoch の学習で十分誤差は減少し、 収束が確認された.分類器の性能評価は 5-fold cross-validation を用いて行った.各分類器の感度(SE: Sensitivity)、特異度(SP: Specificity)、および AUC (ROC カーブの下面積)に関して、実験から得られた結果を、 以下の Table.1 に示す.

	SE (%)	SP (%)	AUC
Training CNN with	84.8	89.5	0.873
Fine-tuning			
Training CNN from	80.9	86.1	0.847
Full-scratch			

Table.1 メラノーマ/母斑の分類性能

以下の Fig. 6 に, 分類結果から得られた ROC カーブを示す.



8. 考察

提案手法である Fine-tuning による転移学習の実施は, ROC カーブにおける AUC による評価で,2.6%の分類性 能を改善した.また,フルスクラッチで学習を行った CNN においても,最終的に示した分類性能は決して悪くはな かったものの,Fine-tuning による分類精度には劣った. さらに,フルスクラッチでの学習は,Fine-tuning による 学習に対し,学習誤差の収束が遅かったという懸念があ る.このことから,転移学習では十分学習能力が備わっ たモデルのパラメータを初期値として用いたことにより, 効率的な学習ができていることを確認した.転移学習は 一般的に事前学習に用いる元ドメインのデータが目標ド メインを内包するものでなければ効果は薄いとされるが, メラノーマ/母斑クラスにも適用できたことは非常に興 味深い結果である.

また、一般的に Fine-tuning はある程度の規模のデータ セットにより行わなければ、十分な知識転移は行えない とされるが、結果では 35,200 枚の小規模なデータセット で、十分に高性能な分類器の構築を行うことができた. 少量の学習データにも関わらず汎用性の高い分類器を構 築できたのは、前処理として行った腫瘍領域の長軸の位 置合わせ処理を実施したことにより、メラノーマや母斑 が持つ形状特徴の差異を強調することができ、学習効果 の高いデータへ変換することができたことに起因すると 考えられる.これにより、メラノーマと母斑を分類する のに適切な特徴量の獲得が行われたのではないかと考える.この前処理の有効性の検討は,以前我々が行った検証の結果より,CNNにおける分類において効果的であることが実証されている [22].

以上の結果から,適切な知識転移を行うことで,元ド メインがカバーしない領域の目標ドメインにおいても, 十分な分類性能を示すことができるという知見を得るこ とができた.

9. 結論

本研究では、転移学習技術を用いた学習済みモデルの Fine-tuning、および医学的知見に基づいた CNN 学習にお ける効果的な前処理の実施により、感度 84.8%、特異度 89.5%、AUC 0.873 の高性能な分類器を構築することがで きた. この数値は熟練した皮膚科専門医に匹敵するもの であり、従来の自動診断システムで実施されてきた「腫 瘍領域抽出」、「特徴量設計」、「特徴量抽出」などの 複雑な処理を伴わずとも、十分な分類性能を達成した. これは、従来の自動診断システムでは扱うことができな かった症例においても診断が可能になったことを意味す る. 今後は診断結果だけでなく、診断結果に至った根拠 の提示が、本研究が次に取り組むべき課題である.

謝辞:本研究の全過程を通して,懇切なる御指導を頂き ました知的情報処理研究室 彌冨 仁 准教授に心より感 謝致します.そして,日頃より本研究に対して多くの御 意見を寄せてくださいました彌冨研究室の皆様に深く御 礼を申し上げます.

参考文献

- 1) F. L. Meyskens Jr. et al. : Natural history and prognostic factors influencing survival in patients with stage I disease, Cancer 1998; Vol.62, Vo.2, pp.1207-1214, 1988.
- 2) H. P. Soyer. et al. : Early diagnosis of malignant melanoma by surface microscopy, Lancet, Vol.2, No.8562, p.803, 1987.
- 3) W. Stolz. et al. : ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma, European Journal of Dermatology, No.4, No.7, pp.521-527, 1994.
- 4) G. Argenziano. et al. : Epiluminescence microscopy for the diagnosis of ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis, Archives of Dermatology, No.134, pp.1536-1570, 1998.
- W. Stolz. et al. : Color Atlas of Dermatoscopy -- 2nd enlarged and completely revised edition, Berlin, Blackwell publishing, 2002.
- 6) M. E. Celebi. et al. : A methodological approach to the classification of dermoscopy images, Computerized Medical Imaging and Graphics, Vol.31, No.6, pp.362-373, 2007.

- 7) H. Iyatomi. 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.
- 8) K. Shimizu. et al. : Four-class classification of skin lesions with task decomposition strategy, IEEE Trans. Biomedical Engineering, Vol.62, No.1, pp.274-283, 2015.
- 9) K. Simonyan. et al. : Very Deep Convolutional Networks for Large-Scale Image Recognition, CoRR, abs/1409.1556, 2014.
- C. Szegedy. et al. : Going Deeper With Convolutions, CoRR, abs/1409.4842, 2014.
- 11) A. Krizhevsky. et al. : ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems, pp.1106-1114, 2012.
- 12) N. Codella. et al. : Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images, in MICCAI MLMI, Vol.9352, pp.118-126, 2015.
- 13) J. Kawahara. et al. : Deep Features to Classify Skin Lesions, IEEE 13th International Symposium on Biomedical Imaging, pp.1397-1400, 2016.
- 14) A. Esteva. et al. : Dermatologist-level classification of skin cancer with deep neural networks, Nature, Vol.542, pp.115-118, 2017.
- 15) S. J. Pan. et al. : A Survey on Transfer Learning, IEEE Transactions on Knowledge and Data Engineering, Vol.22, No.10, pp.1345-1359, 2010.

- 16) A. F. Frangi. et al. : Multiscale vessel enhancement filtering, Medical Image Computing and Computer Assisted Intervention, Vol.1496, pp.130-137, 1998.
- 17) Y. Sato. et al. : Three-dimensional multi-scale line filter for segmentation and visualization of curvatures in medical images, Medical Image Analysis, Vol.2, No.2, pp.143-168, 1998.
- 18) M. Bertalmio. et al. : Navier-stokes, Fluid Dynamics, and Image and Video Inpainting, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.355-362, 2001.
- 19) D. Gutman. et al. : Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC), Proc. IEEE Computer Society Conference Computer Vision and Pattern Recognition, abs/1605.01397, 2016.
- 20) J. Deng. et al. : ImageNet: A large-scale hierarchical image database, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.2-9, 2009.
- 21) Y. Jia. et al. : Caffe: Convolutional Architecture for Fast Feature Embedding, arXiv preprint arXiv: 1408.5093, 2014.
- 22) T. Yoshida. et al. : Simple and Effective Pre-processing for Automated Melanoma Discrimination based on Cytological Findings, Proc. IEEE Big Data 2016 (3rd Big Data Analytic Technology for Bioinformatics and Health Informatics (KDDBHI2016)), pp.3439-3442, 2016.