

皮膚病理画像からダーモスコピー画像切片の色推定

柚木一人[†] 彌富仁[†] 梅垣知子[‡] 田中勝[‡][†]法政大学 理工学部 応用情報工学科[‡]東京女子医科大学東医療センター皮膚科

概要

皮膚腫瘍の診断は、専用の拡大鏡であるダーモスコピーが主であるが診断は主観的で難しい。皮膚病理検査は診断が確定するが侵襲的かつ、メラノーマが対象の場合は転移の恐れがあるため手術前の実施に限られる。病理画像とダーモスコピー画像（D画像）の関係を機械学習で得ることができれば臨床での貢献が大きいが、これらに対応付ける工学的な研究はない。本研究はこれらの関係性を得るための第一歩として、病理画像を入力として、皮膚腫瘍の診断にきわめて重要であるD画像の切片上の色分布を推定する深層学習モデルを試作した。本研究は皮膚科学研究における初めての試みであり、26症例の診断が確定した青色母斑の病理画像からD画像切片の色推定において、交差検定にて平均推定色差6.97を実現した。

1 背景

皮膚がんを含む皮膚腫瘍の診断は肉眼やダーモスコピー画像（D画像）検査により行われるが、診断は熟練の専門医においても診断は難しく、主観的で再現性が乏しい[1]。一方、組織検査は確定診断が得られるが侵襲的であり、また皮膚科領域では転移の可能性があるため、実施は手術直前に限定されるなど制限がある。こうした背景からD画像を対象とした自動診断技術が多く提案されてきた[2][3]。特に深層学習技術とオープンデータセットの活用により研究は活発になり、数字上では熟練皮膚科医を上回る報告も数多くある。しかしながら、それらの診断の根拠の提示や、説明性を満たした研究は極めて限られており、システムの信頼性に課題が残されている[4]。

そのため究極的にはD画像から、組織検査により得

Color Estimation of Dermoscopy Image Section from Skin Pathology Image

KazutoYUNOKI[†], HitoshiIYATOMI[†], NorikoUMEGAKI[‡], MasaruTANAKA[‡],[†]Department of Applied Informatics, Faculty of Science and Engineering, Hosei University
184-8584, Koganei, Japan[‡]Tokyo Women's Medical University Medical Center East
116-8567, Arakawa, Japan

{kazuto.yunoki.7k@stu, iyatomi@,}.hosei.ac.jp

られる病理画像を正しく推定することが可能であれば、皮膚病変の診断にきわめて大きな貢献となる。しかしD画像は皮膚表面からの画像で、病理画像は断面の画像であるため、この推定は典型的な逆問題であり困難である。また、著者らの知る限りこれらの画像を工学的な観点から紐付けた研究は行われておらず、これらの関係性についての技術的な蓄積もない。

本研究はこれらの関係性を得るための第一歩として、病理画像を入力とした皮膚腫瘍の診断にきわめて重要なD画像の切片上の色分布を推定する深層学習モデルを試作し評価した。

2 対象画像

実験では、東京女子医科大学病院から提供された26症例の診断が確定した青色母斑に対する病理画像および、切片部位が記載された対応するD画像（各患者から1ペアずつ）を用いた。

2.1 病理画像データセット

使用した症例の病理画像例を図1に示す。病理画像はヘマトキシリン・エジオン（HE）染色し、組織標本としてとても高い解像度で診断に用いられている。表皮から約1mmまでの組織構造がD画像に反映され、メラニン、血管などの皮膚組織が皮膚表面の色に影響する。茶色に見える部分がシミや母斑の原因となるメラニン色素であり、浅い部分に多く存在するとD画像上で茶色に見え、深い部分に多く存在すると青黒く見える。本研究では、表皮から1mmが収まるように病理画像をリサイズし、224×224の画像として使用した。

2.2 ダーモスコピー画像データセット

使用した症例のD画像例を図2に示す。D画像は病理画像とセットで与えられ、専門医が切除した部分が長方形の青いボックスで示されているが、厳密な位置は明確ではない。そこで、本研究では長方形の短軸方向に平均化して1×196の色情報を取得し教師データとした。



図 1: 病理画像



図 2: ダーモスコピー画像



図 3: D 画像色分布の予測結果 (上:教師 下:予測)

3 方法

3.1 ネットワークモデルの構成

病理画像から特徴量を抽出して D 画像の色分布の推定を行うにあたり、病理画像との位置情報の保持が重要である。そのため本研究では、畳み込み層、バッチノーマライゼーション、活性化関数を順に重ね合わせた畳み込み層のみの Fully Convolutional Network (FCN) を独自に改良した 5 層 FCN を構築した。入力画像は前述の通り 224×224 , RGB3 チャンネルの病理画像、出力は対応する D 画像切片的の 196 画素の RGB3 チャンネルである。

3.2 データセットの前処理

本実験では利用可能な画像データが極めて限られるため、学習データに対して位置情報の保持を考慮し以下の計 $6 \times 2 = 12$ 倍のデータの水増しを行った。(1) 病理画像の HE 染色での染まり具合はばらつきがあるため、計 5 種類のガンマ値 (0.55, 0.70, 0.85, 1.25, 1.50) で変換。(2) 病理画像と D 画像それぞれに対して水平反転。

3.3 評価方法

学習と評価には 5-fold cross validation を用いた。本研究では D 画像の切片の色推定の評価に、 L^*a^*b 表色系を用い、検証データの推定画素の各ピクセルの平均色差を算出した。

4 結果

検証データに対する D 画像切片の色分布の予測結果の例を図 3 に示す。この例の推定色差は 3.77, 4.25, 13.4 であり、全体のダーモスコピー切片の色の推定誤差は、色差 6.97 ± 1.45 であった。一部の結果には図 3 下段の例のような周期的なノイズが表れ、切片の推定精度を低下させた。これは図 3 上段のような良好な結果においても、腫瘍部分以外のところで観察された。

5 考察と今後の課題

推定結果の周期的なノイズは、他の症例でも同様に表れているが、これは学習データの横方向に対する位

置合わせ不足であると考えられる。症例によって腫瘍の位置、大きさがばらばらであるため、腫瘍ではない位置に腫瘍の色のようなスリットがノイズとして表れていると言える。今後は腫瘍の横方向に対する位置合わせをも検証する必要があると考えられる。

日本工業規格によると色差が 3.2~6.5 は色彩管理で一般的に扱われる許容色差であり、我々の実験では一般的には同色とされるレベルに準じる程度で病理画像から D 画像切片の色分布の予測が実現できたと言える。

本研究では学習データが極めて限られるため、深層学習モデルで扱いやすい低解像度 (224×224) でモデルを実装しているが、今後より多様なデータを収集し、より高解像度の画像が扱えるネットワークモデルを構築し、検証する必要があると考えられる。

6 結論

病理画像から D 画像切片の色を一般的には同色と見なせるレベルまで予測できる深層学習モデルを作成することができた。

参考文献

- [1] G. Argenziano, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, M. Binder, L. Cerroni, G. De Rosa, G. Ferrara *et al.*, "Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet," *Journal of the American 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, and K. Ogawa, "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] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, and A. Schlaefer, "Skin lesion classification using ensembles of multi-resolution efficientnets with meta data," *MethodsX*, p. 100864, 2020.
- [4] S. Murabayashi and H. Iyatomi, "Towards explainable melanoma diagnosis: Prediction of clinical indicators using semi-supervised and multi-task learning," in *2019 IEEE International Conference on Big Data (Big Data)*. IEEE, 2019, pp. 4853–4857.