

頑健な皮膚腫瘍診断支援のための body hair augmentation

北田 俊輔 彌富 仁

法政大学 大学院 理工学研究科 応用情報工学専攻

概要

悪性度の高い皮膚がんの1つであるメラノーマは早期発見が重要である。本研究では頑健な自動診断システム実現のため、深層学習分野で極めて優れた成果を実現している要素技術である residual networks, squeeze-and-excitation networks, mean teachers を導入した識別器をベースに、皮膚腫瘍診断のための独自の body hair augmentation を提案する。深層学習技術に基づく識別器に対して提案手法を適用することにより、識別精度 90.6% を達成した。

1 はじめに

メラノーマは予後不良の皮膚がんであり早期発見が必要不可欠である。診断時は医師の目視検査やダーモスコピー (皮膚科医専用の偏光レンズつき拡大鏡) による診断を行うが、診断精度は 75-84% 程度 [1] に留まる。

近年, convolutional neural networks (CNN) を皮膚腫瘍画像解析に利用した高精度の自動診断支援手法が提案されてきている [2]。このような CNN ベースの診断システムは学習に際して大量の訓練データが必要であるが、医療データの大規模データセット構築は容易ではない。こうした問題に対して ImageNet で事前学習した深層学習モデルを fine-tuning することにより、診断精度の向上が報告されている [3]。しかしながら皮膚ダーモスコピー画像は非常に多様な外観を持つことから過学習に陥りやすい。また体毛や影、コントラスト不良など様々な悪条件画像に対しては大幅に診断精度が低下し、頑健な識別器の構築は容易ではない。

一方、画像認識で最先端の結果を残している SE-ResNet [4] は ResNet [5] に対して空間特徴を圧縮して伝達する squeeze-and-excitation block (SEBlock) を導入し、結合加重の総数を減らすことで過学習を抑制し、高い識別精度を実現している。また複数の識別器を組み合わせる network ensemble と半教師あり学習

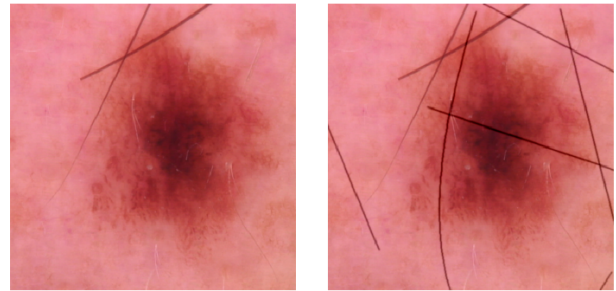


図 1: Body hair augmentation 適用前 (左図) と適用後 (右図)。皮膚腫瘍画像に対して疑似体毛を付与する。

を同時に成り立たせる mean teachers [6] は、汎化性向上に大きな貢献をもたらす手法で、今回のように教師 (診断) ラベルのコストの高い医療分野での貢献が特に期待できる。我々はこの SE-ResNet と mean teachers を組み合わせた高い頑健性が期待できるネットワークをメラノーマ診断のためのベースラインとして用いる。

先行研究ではダーモスコピー上の体毛といったノイズは前処理で取り除かれてきた [7]。しかしこうした前処理では、もともと体毛が存在していた皮膚部分や腫瘍領域は擬似的なものになってしまう。そこで本研究では、悪条件画像においても安定的に識別ができることを目的に、患部付近の体毛がシステムの診断の妨げとなる点に着目し、あえて疑似体毛を付与することで頑健となるよう学習を行う body hair augmentation を提案する。前述の深層学習技術に基づく複数種の識別器と半教師あり学習の枠組みに対して提案手法を適用することで効果を確認する。

2 手法

本報告で提案する body hair augmentation は buffon の針 [8] に基づいて、擬似的に体毛を模倣する線を描画する。このとき描画される線はベジェ曲線に従ってランダムに曲線とする。描画した疑似体毛はガウシアンフィルタを掛けた後、差分合成を用いてダーモスコピー画像に合成する。提案手法の適用例を図 1 に示す。

表 1: ISIC2018 公式評価セットを用いた皮膚腫瘍画像に対する識別精度の比較

Model	Balanced Accuracy [%]			
	Baseline	Mean teachers	Body hair augmentation	Mean teachers + Body hair augmentation
SE-ResNet101 [4]	86.4	87.3	88.4	90.6
ResNet101 [5]	85.7	86.9	87.5	88.5

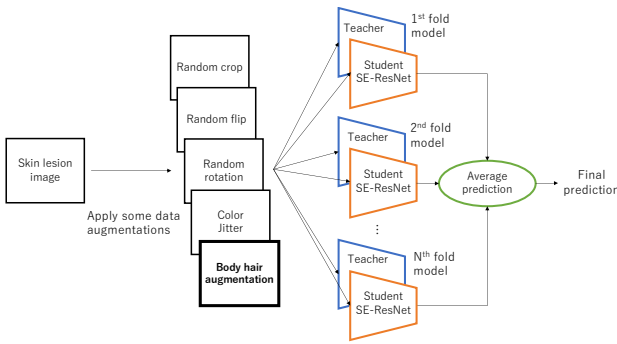


図 2: 本研究で用いたシステムアーキテクチャの概要

3 評価実験と結果

評価実験では表 2 に示すメラノーマ (MEL) を含む 7 種類の皮膚腫瘍画像が含まれる, HAM10000 [9] をベースとした ISIC2018 Challenge¹task 3 データセットを学習データとして使用した。

表 2: ISIC2018 task 3 データセットの概要

	labeled						unlabeled		Total	
	MEL	NV	BCC	AKIEC	BKL	DF	VASC	validation		test
	1,113	6,705	514	327	1,099	115	142	193	1,512	11,720

同じ患者の皮膚画像が学習セットと評価セットに含まれないように分割する, group 5-fold cross validation を用いて 5 つのモデルを学習させた。システムは ResNet101 および SE-ResNet101 を用いたネットワークに対して, 半教師あり学習である mean teachers を用いたアーキテクチャ [10] をベースとしている。システムの概要を図 2 に示す。学習時には入力画像に data augmentation として random crop, flip, rotation, color jitter に加えて, 提案手法である body hair augmentation を適用した。Mean teachers のハイパーパラメータとして, ema decay rate=0.999, consistency loss に mean squared error を使用した。予測時には cross validation で学習させた 5 つのモデルの予測値を平均し, 最終的な予測結果とした。

ISIC2018 公式評価セットに対するクラスごとの平均

¹ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection <https://challenge2018.isic-archive.com/>

識別率の平均 (balanced accuracy) の結果を表 1 に示す。ベースラインのモデルに対して半教師あり学習の mean teachers を適用することにより 1% 程度の精度向上が見られた。また提案手法を適用することにより 3% 程度の精度向上が確認できた。最終的に半教師あり学習と提案手法を組み合わせることで, 平均して 4% 程度の識別精度向上を達成した。

4 おわりに

本研究では, 体毛の有無に頑健となるように学習を行う body hair augmentation を導入した。最先端の深層学習モデルに対して提案手法を適用することで診断精度およびモデルの頑健性向上を確認した。

参考文献

- [1] M. Vestergaard, P. Macaskill, P. Holt, and S. Menzies, "Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting," *British Journal of Dermatology*, vol. 159, no. 3, pp. 669–676, 2008.
- [2] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, p. 115, 2017.
- [3] A. Menegola, M. Fornaciali, R. Pires, S. Avila, and E. Valle, "Towards automated melanoma screening: Exploring transfer learning schemes," *CoRR arXiv:1609.01228*, 2016.
- [4] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *IEEE Proc. CVPR*, 2018.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *IEEE proc. CVPR*, pp. 770–778, 2016.
- [6] A. Tarvainen and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," *Advances in neural information processing systems*, pp. 1195–1204, 2017.
- [7] T. Lee, V. Ng, R. Gallagher, A. Coldman, and D. McLean, "Dullrazor®: A software approach to hair removal from images," *Computers in biology and medicine*, vol. 27, no. 6, pp. 533–543, 1997.
- [8] G. L. L. De Buffon, "Essai d'arithmétique morale," *Euvres philosophiques*, 1777.
- [9] P. Tschandl, C. Rosendahl, and H. Kittler, "The ham10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions," *CoRR arXiv:1803.10417*, 2018.
- [10] S. Kitada and H. Iyatomi, "Skin lesion classification with ensemble of squeeze-and-excitation networks and semi-supervised learning," *CoRR arXiv:1809.02568*, 2018.