ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/375074151

A Machine Learning Framework for Hair Type Categorization to Optimize the Hair Removal Algorithm in Dermatoscopy Images

Conference Paper · December 2023

DOI: 10.1109/IEEECONF58974.2023.10404510

CITATIONS
3

READS
269

7 authors, including:

Georgios S. Ioannidis
Foundation for Research and Technology Hellas
28 PUBLICATIONS **286** CITATIONS

SEE PROFILE

Eleftherios Trivizakis
Foundation for Research and Technology Hellas
51 PUBLICATIONS **702** CITATIONS

SEE PROFILE

Aimilios Lallas
Aristotle University of Thessaloniki
578 PUBLICATIONS **14,890** CITATIONS

SEE PROFILE

George Evangelou
University of Crete
55 PUBLICATIONS **496** CITATIONS

SEE PROFILE

All content following this page was uploaded by Eleftherios Trivizakis on 08 December 2023.

The user has requested enhancement of the downloaded file.

A Machine Learning Framework for Hair Type Categorization to Optimize the Hair Removal Algorithm in Dermatoscopy Images

Georgios S. Ioannidis, Eleftherios Trivizakis, Konstantinos Krasagakis, Aimilios Lallas, Zoe Apalla, Georgios Evangelou and Kostas Marias

***Abstract*— This work proposes a machine learning (ML) framework to classify the hair type of dermatoscopy images into four classes using the imaging features taken from the binary hair contour masks. Furthermore, the optimal kernel of the black-hat hair removal algorithm is then examined through the structural similarity index measure (SSIM) between the original and the pre-processed image. The best performance of the classification model in terms of ACC and AUC was obtained by the SVM classifier, achieving 80% and 79.8%, respectively. A kernel size of up to 20 by 20 is proposed for image filtering without significant loss of texture information in the lesion.**

***Clinical Relevance*— This paper presents an effort to find the optimal kernel size of the black-hat algorithm for hair removal on dermatoscopy images, while maintaining high image quality. This work is proposed as an automated pre-processing step for deep learning in skin disease classification.**

I. Introduction

Digital hair removal from dermatoscopy images is a crucial preprocessing step in the field of dermatology. Dermatoscopy is a non-invasive imaging technique used for the examination of skin lesions to aid in the early detection of melanoma and other skin cancers. However, these images often contain unwanted hair, which can obscure the visualization of the underlying features and make accurate analysis challenging, especially for deep learning models in skin disease classification [1]. To address this issue, digital hair removal techniques have been developed to enhance the utility of dermatoscopy as a diagnostic tool and contribute to the early detection and treatment of skin diseases. One common approach to digital hair removal involves image processing algorithms that detect and remove hair artifacts while preserving the integrity of the skin lesion. These algorithms typically rely on color and texture analysis to distinguish between hair and skin pixels, followed by the application of various filters and morphological operations such as the black-hat algorithm to remove the detected hair structures [2]. However, the results of this algorithm depend on the size of the kernel used, image dimensions, and the hair type on the skin (thin, medium, and thick) as can be seen in Fig. 1. To the best of our knowledge, the proposed analysis is the first ML framework to assess hair thickness and the impact of hair removal algorithms on the region of interest. Additionally, a method for identifying the optimal kernel size for the Black-hat algorithm is presented.

II. Methods

In order to explore the optimal kernel of the Black-hat algorithm, 600 images from the ISIC 2017 dataset [3] were used and labeled into four classes (no hair, thin, medium, and thick hair). The number of images in each class was 423, 29, 35, and 113 for no hair, thin hair, medium hair, and thick hair, respectively. This work constitutes a crucial pre-processing step for downstream tasks such as dermatoscopy image lesion detection and classification with deep learning. Therefore, the aforementioned images were resized to 640 x 640 pixels to accelerate the hair removal algorithm but also to be compatible with the state-of-the-art detection deep models. For the classification of each image, 1121 features were extracted with the pyradiomics library [4], including features such as first- and higher-order statistics, texture features such as grey-level run length matrix, grey-level co-occurrence matrix, grey-level size zone matrix, grey-level difference matrix, and shape-based 2D features. Additionally, local binary patterns and image transformation methods such as Laplace of Gaussian, logarithmic, exponential, and gradient were also used. The masks used to obtain the imaging features are the binary image contours of the black-hat algorithm with a 20 by 20 kernel. The differentiation between hair types was performed using a variety of classifiers, including the support vector machine classifier with the radial basis function kernel, which has been widely used in medical image classification problems [5]–[7], the quadratic discriminant analysis, random forest classifier, logistic regression, adaboost, and k-nearest neighbors (k=4) classifiers. The classifiers were trained in a 3-fold cross-validation scheme using the 1121 imaging features. To avoid sample selection bias and overfitted

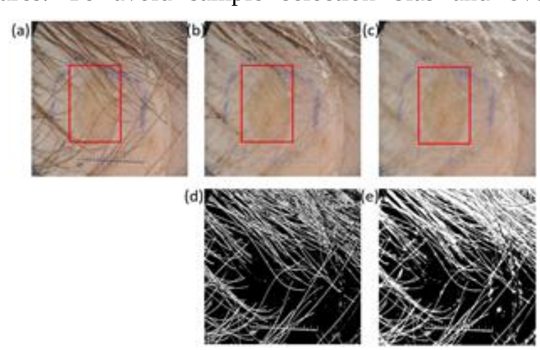

Figure 1. Examples of Black-hat hair removal algorithm. (a) original image, (b), (d) removed hair with kernel = (10, 10) and hair contours mask respectively. (c), (e) removed hair with kernel = (30, 30) and hair contours mask respectively.

G.S.I., E.T., and K.M. are with the Computational Biomedicine Laboratory of Foundation for Research and Technology Hellas, Heraklion, Greece; K.K. and G.E. are with the Department of Dermatology, University Hospital of Heraklion, Heraklion, Greece. A.L. is with the First Department of Dermatology, School of Medicine, Faculty of Health Sciences, Aristotle University, Thessaloniki, Greece. Z.A. is with the Second Dermatology Department, Aristotle University of Thessaloniki, Greece.

E.T. is the corresponding author, e-mail: trivizakis@ics.forth.gr

models, the data stratification was applied on an image basis with respect to the class representation across folds. To overcome the imbalanced number of samples in each class, the synthetic minority oversampling technique (SMOTE) was applied to the training data, and the trained models were evaluated exclusively on the unseen testing sets [8]. Furthermore, a three-step feature selection process was used to reduce the feature space. Firstly, a variance threshold was applied to remove features with zero variance. Secondly, a univariate method (ANOVA, analysis of variance) was used to remove noisy information from the feature space. The last step was to apply the multivariate method of logistic regression based on the L1 penalty, which minimizes the coefficients of the redundant features. Hair removal is then followed by these steps: a) conversion of the original image into grayscale, b) selection of the kernel and application of the morphological black-hat transformation on the grayscale image to retrieve the hair contours (i.e. images (d) and (e) in Fig. 1), c) binarization of the contour image, and d) application of the inpainting algorithm on the original image using the mask of step c). The OpenCV implementation of black-hat and inpaint algorithms was used. Lastly, the evaluation of the processed images after hair removal is achieved by the SSIM, which ranges from 0 to 1. The SSIM was calculated between the original and the corresponding images with the hair removed using the bounding boxes (red rectangle in Fig. 1) that include the underlying skin pathology. These bounding boxes were obtained by a deep learning-based detection model (large YOLOv8 [9]) trained on the ISIC 2018, which includes annotated lesion regions.

III. RESULTS

The performance metrics, accuracy (ACC) ± standard deviation (std) and the area under the curve (AUC) ± std of the hair type classification scheme are presented in Table I.

TABLE I. TERNARY CLASSIFICATION PERFORMANCE

Model evaluation Metrics		
Classifier	***ACC ± std***	***AUC ± std***
Quadratic Discriminant Analysis	80.3 ± 2.7	66.7 ± 1.9
Support Vector Machine	80.0 ± 0.7	79.8 ± 2.8
Random Forest Classifier	76.5 ± 1.4	79.3 ± 2.7
Logistic Regression	71.5 ± 2.0	79.6 ± 3.2
Adaboost Classifier	64.7 ± 4.1	71.7 ± 2.5
K-Nearest Neighbors Classifier	62.5 ± 0.7	71.3 ± 1.5

The results of Table II in conjunction with visual inspection of the preprocessed images were used to optimize the kernel size of the Black-hat algorithm, achieving minimal image quality loss (approximately 5% of SSIM) with the best removal of hair. The mean SSIMs for each image type are summarized in Table II with respect to the kernel used.

TABLE II. THE MEAN STUCTURAL SIMILARITY INDEX MEASURE ACROSS LESION BOUNDING BOXES WITH DIFFERENT TYPE OF HAIR

	SSI per hair type			
Kernel size	***Thin***	***Medium***	***Thick***	***No hair***
5 by 5	0.985	0.972	0.966	0.992
10 by 10	0.966	0.955	0.933	0.979
15 by 15	0.950	0.941	0.917	0.966
20 by 20	0.914	0.922	0.898	0.946
30 by 30	0.878	0.883	0.863	0.899
40 by 40	0.843	0.839	0.830	0.839

IV. DISCUSSION & CONCLUSION

In this work, a ML framework was implemented to identify the hair type of dermatoscopy images into four classes using imaging features taken from the binary hair contour masks. The highest performance was achieved by the SVM classifier, yielding an ACC of 80% and an AUC of 79.8%, as shown in Table I. Additionally, the black-hat algorithm was used on the lesion images containing hair and on samples without hair with a variable kernel size in order to identify the optimal size for each hair type and assess potential alterations in the texture of the lesion. The SSIM was calculated over the region of interest (ROI) to assess the impact of the hair removal algorithm. A control set of samples with no hair was used to quantify the changes in texture and identify the trade-off between removing hair and introducing noise into the lesion region. A difference of less than 6% was observed with kernel sizes between 5 by 5 and 20 by 20. Alterations greater than 10% for the large kernels of 30 by 30 and 40 by 40 can also be observed. In particular, the difference between the group with hair and the control group for the large kernels reflects the extensive alterations in terms of lesion texture and not just the results of the hair removal process. Thus, a kernel size of up to 20 by 20 is proposed for image filtering without significant loss of texture information about the lesion.

REFERENCES

[1] P. Ghosh *et al.*, "SkinNet-16: A deep learning approach to identify benign and malignant skin lesions," *Front Oncol*, vol. 12, p. 931141, Aug. 2022, doi: 10.3389/FONC.2022.931141/BIBTEX.

[2] A. H. Khan, D. N. F. Awang Iskandar, J. F. Al-Asad, and S. El-Nakla, "Classification of Skin Lesion with Hair and Artifacts Removal using Black-hat Morphology and Total Variation," *International Journal of Computing and Digital Systems*, vol. 10, no. 1, pp. 2210–142, 2021, doi: 10.12785/ijcds/100157.

[3] N. C. F. Codella *et al.*, "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)," *CoRR*, vol. abs/1710.05006, 2017, [Online]. Available: http://arxiv.org/abs/1710.05006

[4] J. J. M. Van Griethuysen *et al.*, "Computational radiomics system to decode the radiographic phenotype," *Cancer Res*, vol. 77, no. 21, pp. e104–e107, Nov. 2017, doi: 10.1158/0008-5472.CAN-17-0339.

[5] E. Trivizakis *et al.*, "A novel deep learning architecture outperforming 'off-the-shelf' transfer learning and feature-based methods in the automated assessment of mammographic breast density," *Oncol Rep*, Sep. 2019, doi: 10.3892/or.2019.7312.

[6] G. S. Ioannidis *et al.*, "Quantification and Classification of Contrast Enhanced Ultrasound Breast Cancer Data: A Preliminary Study," *Diagnostics 2022, Vol. 12, Page 425*, vol. 12, no. 2, p. 425, Feb. 2022, doi: 10.3390/DIAGNOSTICS12020425.

[7] G. S. Ioannidis, E. Trivizakis, I. Metzakis, S. Papagiannakis, E. Lagoudaki, and K. Marias, "Pathomics and Deep Learning Classification of a Heterogeneous Fluorescence Histology Image Dataset," Applied Sciences, vol. 11, no. 9, p. 3796, Apr. 2021, doi: 10.3390/app11093796.

[8] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.

[9] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLOv8." 2023. [Online]. Available: https://github.com/ultralytics/ultralytics

View publication stats