

# Bias Analysis and Data Augmentation Strategies in Skin Lesion Classification

Análise de Viés e Estratégias de Aumento de Dados na Classificação de Lesões de Pele

Thiago Meneses Lopes<sup>1\*</sup>, Martin Heckmann<sup>1</sup>, Marcelo Ponciano da Silva<sup>2</sup>, Pedro Henrique Bugatti<sup>1</sup>, Priscila Tiemi Maeda Saito<sup>1</sup>

**Abstract:** Early detection of skin cancer is vital for effective treatment and improving patient recovery. In recent years, a growing number of computer vision studies have been developed to aid in diagnosis, drawing significant attention from researchers. However, challenges still persist, such as data imbalance and the lack of comprehensive datasets. Additionally, limited research has focused on how variations in skin tone across different populations affect the performance of models in skin lesion classification. This study seeks to create a more effective approach to address data biases in lesion classification across diverse skin tones. We initially explored several data augmentation techniques, employing traditional feature extractors for image analysis. For classification, models such as k-Nearest Neighbor, Random Forest, and Support Vector Machine were used. This study focused on two well-known and publicly available skin lesion datasets: HAM10000 and PAD-UFES-20, both of which have significant class imbalances. Further experiments were conducted to assess potential biases, with the Individual Typology Angle (ITA) metric applied to evaluate the skin tone distribution within the datasets.

**Keywords:** data augmentation — classification — skin lesion — computer vision — data bias

**Resumo:** A detecção precoce do câncer de pele é essencial para um tratamento eficaz e para a melhora na recuperação dos pacientes. Nos últimos anos, um número crescente de estudos em visão computacional tem sido desenvolvido para auxiliar no diagnóstico, atraindo grande atenção da comunidade científica. No entanto, ainda persistem desafios, como o desbalanceamento dos dados e a ausência de conjuntos de dados abrangentes. Além disso, poucas pesquisas têm investigado como as variações de tom de pele entre diferentes populações afetam o desempenho dos modelos na classificação de lesões de pele. Este estudo busca desenvolver uma abordagem mais eficaz para lidar com vieses de dados na classificação de lesões em diferentes tons de pele. Inicialmente, foram exploradas diversas técnicas de aumento de dados, empregando extratores de características tradicionais para análise das imagens. Para a classificação, foram utilizados modelos como k-Nearest Neighbor (k-NN), Random Forest (RF) e Support Vector Machine (SVM). O estudo concentrou-se em dois conjuntos de dados de lesões de pele amplamente conhecidos e publicamente disponíveis — HAM10000 e PAD-UFES-20 —, ambos apresentando significativo desbalanceamento entre classes. Experimentos adicionais foram conduzidos para avaliar possíveis vieses, aplicando-se a métrica Individual Typology Angle (ITA) para examinar a distribuição dos tons de pele dentro dos conjuntos de dados.

**Palavras-Chave:** aumento de dados — classificação — lesão de pele — Visão computacional — Viés de dados

<sup>1</sup> Departamento de Computação, Universidade Federal de São Carlos (UFSCar), Brazil

<sup>2</sup> Instituto Federal do Triângulo Mineiro (IFTM), Brazil

\*Corresponding author: thiagoml@estudante.ufscar.br

DOI: <http://dx.doi.org/10.22456/2175-2745.150974> • Received: 15/10/2025 • Accepted: 03/11/2025

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

## 1. Introduction

The incidence of skin cancer has been rising each year, primarily due to skin damage caused by ultraviolet radiation. Dermoscopy, a non-invasive diagnostic technique, enables the visualization of morphological features that cannot be seen with the naked eye. While it has enhanced diagnostic performance,

studies [1, 2, 3] have shown that dermoscopy's diagnostic accuracy is lower when performed by novice dermatologists compared to those with more experience. Additionally, even for highly trained professionals, lesion identification largely depends on visual inspection, a process that remains highly subjective.

Despite ongoing efforts in the literature, many studies have notable limitations. For instance, small and imbalanced datasets often include only a limited number of representative samples for each type of skin lesion and various skin tones [4], which can lead to biases—whether data-related, algorithmic, or racial. To address some of these challenges, researchers have explored data augmentation techniques [5, 6, 7, 8, 9]. However, there has been limited attention given to tackling bias in existing studies [10, 11, 12, 13]. Consequently, there is significant potential to enhance this area by developing more diverse datasets and robust learning strategies that are both effective and efficient in identifying different lesion types and skin tones.

Our contributions in this work are: (i) the exploration of data augmentation scenarios in the context of skin lesions, (ii) evaluation of the best combination of feature extractor and classifier, (iii) analysis of the impact of different data augmentation techniques on the feature extraction process, (iv) the categorization of skin tones within skin lesion datasets, and (v) the analysis of how imbalanced skin tone distribution affects the skin lesion classification process.

## 2. Related concepts

### 2.1 Skin tone categorization

The categorization of skin tones in dermatology still lacks consensus. The Fitzpatrick Skin Type (FST) method is widely used and classifies skin into six categories based on its response to ultraviolet light. However, it is a subjective method, determined by observing the skin's reaction to the sun.

On the other hand, the Individual Typology Angle (ITA) method offers an objective approach, using metrics from the CIELab color space to calculate skin tone. Details of the skin tone categorization process are presented in Section 3.4.

### 2.2 Skin lesions and cancer

Skin cancer occurs due to mutations in the DNA of cells, leading to their uncontrolled proliferation. There are two main types: non-melanoma, which is more common and less risky, and melanoma, which is more severe due to its potential for metastasis. Early diagnosis is crucial and can be aided by the ABCDE rule, which evaluates asymmetry, borders, color, diameter, and evolution of lesions.

Non-invasive methods, such as dermoscopy, allow for a more detailed analysis of lesions, complementing invasive exams like biopsies. Additionally, automated systems based on medical imaging and machine learning have shown promise in assisting diagnosis. However, challenges such as data imbalance and the scarcity of labeled images still hinder progress.

## 3. Proposed Methodology

This section outlines our approach for bias analysis and data augmentation strategies in skin lesion classification. Figure 1 illustrates the methodology used in this process.

From collected dermoscopic images, a data augmentation step is performed by applying various visual transformations (scenarios labeled A to G) to enhance sample variability and improve the model's robustness. The method simultaneously categorizes skin tones using the ITA (Individual Typology Angle) index. It involves resizing images and extracting patches from different lesion regions, including structured, central, and random areas. This approach subsequently organizes the data into six skin tone categories, from lighter tones (Type 1) to darker tones (Type 6). Based on these categories, the images undergo feature extraction and classification. Finally, a bias analysis is conducted to assess the classifier's performance across different skin tone subgroups, to identify potential disparities related to data representation and model performance. Our proposed method focuses on exploring data augmentation within the context of skin lesion classification, while also analyzing potential data and algorithmic biases.

### 3.1 Datasets

To evaluate our proposed methodology, we used the HAM10000 and PAD-UFES-20 datasets. The HAM10000 dataset [14] contains 10,015 dermoscopic images from the Department of Dermatology at the Medical University of Vienna. It includes seven diagnostic categories: melanocytic nevus (Nev), melanoma (Mel), basal cell carcinoma (Bcc), actinic keratosis (Akiec), benign keratosis-like lesions (SK), dermatofibroma (Df), and vascular lesions (Vasc). The sample distribution across these categories is uneven, with one class accounting for over 60% of the total dataset. Although it includes a wide range of skin lesion classes, it does not take into account variations in skin tones according to the Fitzpatrick scale.

The PAD-UFES-20 dataset [15], on the other hand, includes patient data and clinical images captured with smartphones. It contains information from 1,373 patients, featuring 1,641 skin lesions and 2,298 images across six diagnostic categories, including three skin diseases and three types of skin cancer. Notably, 58.4% of the lesions in this dataset are biopsy-proven, including all of the identified skin cancers. Although it incorporates different skin tones (based on the Fitzpatrick scale) and several classes of skin lesions (Nev, Sk, Bcc, Akiec, Mel and Scc), it presents a limited number of samples for each class.

Table 1 presents the number of images in each class, highlighting the imbalance in sample distribution across different lesion types. This imbalance becomes even more pronounced when considering the distribution of samples across both skin tone categories and lesion types simultaneously.

### 3.2 Data Augmentation

Multiple image transformation techniques, such as geometric and color-based augmentations, were systematically tested in the data augmentation scenarios. Importantly, the class imbalance in the number of samples was carefully maintained to evaluate the impact of each augmentation technique, ensuring the reliability of our findings.

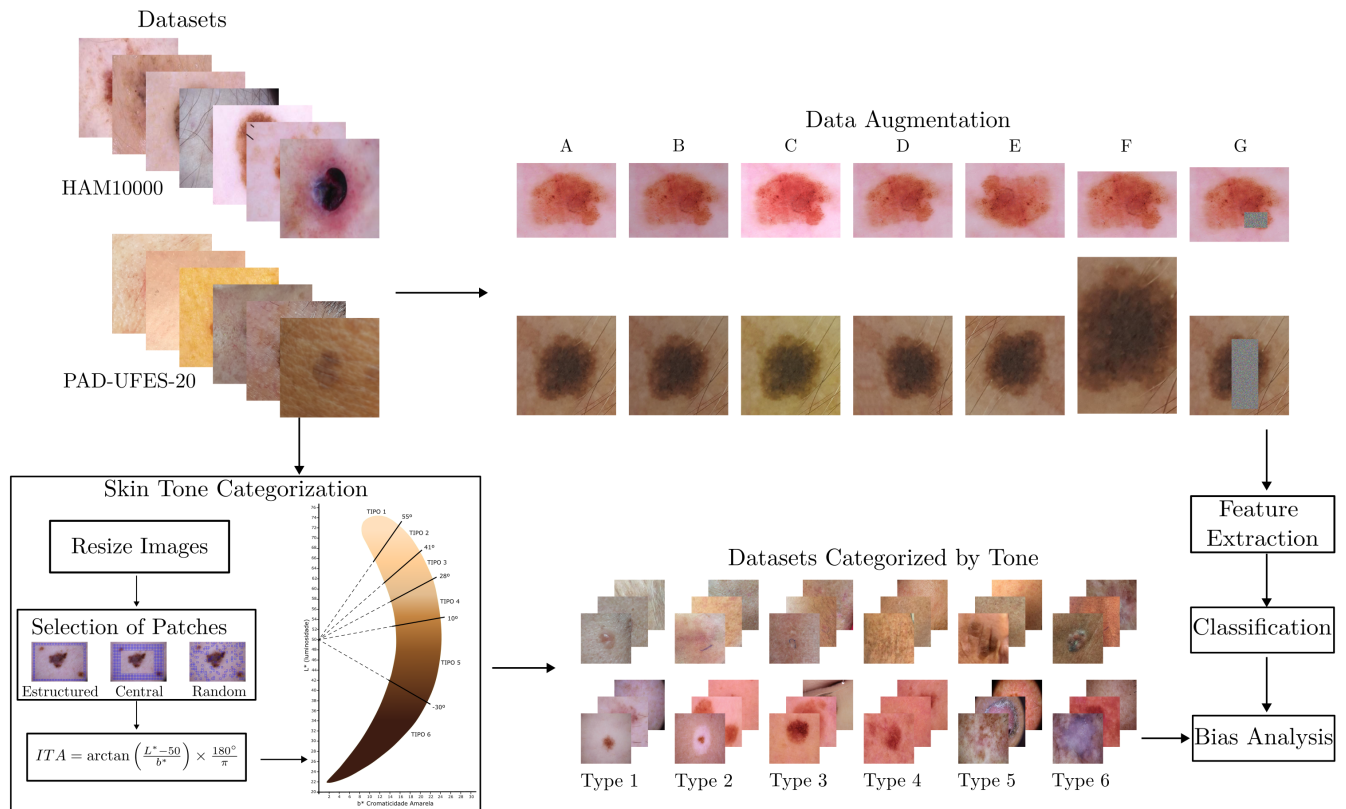


Figure 1. Pipeline of the proposed approach

Table 1. Description of the datasets

| Dataset            | Fitzpatrick | Images | Nev  | Sk   | Bcc | Akiec | Df  | Vasc | Mel  | Sc  |
|--------------------|-------------|--------|------|------|-----|-------|-----|------|------|-----|
| ISIC-18 (HAM10000) | No          | 10015  | 6705 | 1099 | 514 | 327   | 115 | 142  | 1113 | –   |
| PAD-UFES-20        | Yes         | 2298   | 244  | 235  | 845 | 730   | –   | –    | 52   | 192 |

Table 2. Scenarios considered in the experiments include different data augmentation techniques and data imbalance

| ID | Name                                           | Description                                                                                                                                           |
|----|------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| A  | No Augmentation                                | No data augmentation. Only preprocess images.                                                                                                         |
| B  | Saturation, Contrast and Brightness            | Modify saturation, contrast, and brightness by random factors sampled from an uniform distribution of $[0.2, 1.3]$ .                                  |
| C  | Saturation, Contrast, Brightness, and Hue      | As described in B, but also shift the hue by a value sampled from an uniform distribution of $[-0.1, 0.1]$ .                                          |
| D  | Affine                                         | Rotate the image by up to $90^\circ$ , shear by up to $20^\circ$ , and scale the area by $[0.8, 1.2]$ . New pixels are filled symmetrically at edges. |
| E  | Flips                                          | Randomly flip the images horizontally and/or vertically.                                                                                              |
| F  | Random Crops                                   | Randomly crop the original image. The crop has 0.4 - 1.0 of the original area, and 3/4 - 4/3 of the original aspect ratio.                            |
| G  | Random Erasing                                 | Random occlusion in the image. Applies up to 30% occlusion to the image with a 50% de chance.                                                         |
| H  | All augmentation techniques (scenarios B to G) | B + C + D + E + F + G<br>No inclusion of scenario A (i.e. original images)                                                                            |
| I  | All augmentation techniques (scenarios A to G) | A + B + C + D + E + F + G<br>Including scenario A (i.e. original images)                                                                              |

Scenario A consists solely of the original images, while in scenarios B to H, only the samples generated by applying the respective augmentation technique are included. Table

2 outlines the explored scenarios and describes each data augmentation technique, adapted from [12].

In addition to scenarios that apply each augmentation

technique individually, two additional scenarios (H and I) are presented, which involve applying all augmentation techniques (scenarios B to G) with and without the inclusion of the original images, respectively. After applying data augmentation, relevant features are extracted from the images to enhance analysis and classification.

### 3.3 Feature Extraction and Classification

Feature extraction was performed using traditional extractors based on color, shape, or texture. Table 3 provides a list of the extractors, along with their respective categories and dimensionalities. From the extracted features, classifiers are trained and used to classify the types of skin lesions.

For the classification task, we employed three different models: K-Nearest Neighbors (KNN), Random Forest (RF), and Support Vector Machine (SVM). We configured the KNN classifier with five neighbors and employed a distance-based weighting scheme, which assigns greater influence to closer neighbors during the classification process. To build the RF model, we used 200 decision trees, limited the maximum depth to 20, and set constraints requiring at least five samples to split a node and two samples per leaf. We also applied class-balanced weighting and set a fixed random seed to ensure reproducibility. For the SVM classifier, we performed a randomized search over a grid of hyperparameters, testing different kernel types (linear, RBF, polynomial, and sigmoid), regularization values ( $C = 0.01$  to  $100$ ), and gamma values (including 'scale', 'auto', and specific numeric values). We selected the best-performing configuration found during this search for the final evaluation. To estimate the performance and stability of each model on the training set, we applied 5-fold cross-validation, computing the average accuracy and its standard deviation across the folds.

**Table 3.** Feature Extractors for Describing and Classifying Skin Lesions and Tones

| Extractor | Category | Features |
|-----------|----------|----------|
| BIC       | Color    | 128      |
| FOM       | Texture  | 32       |
| GCH       | Color    | 30       |
| Haralick  | Texture  | 13       |
| LBP       | Texture  | 256      |
| TAS       | Texture  | 162      |
| Zernike   | Shape    | 72       |

To analyze bias and assess its impact on classifier performance, the datasets must be evaluated. If they are not already annotated and organized by skin tone, a skin tone categorization process must be applied.

### 3.4 Skin Tone Categorization

Initially, for skin tone categorization, images need to be resized. Then, the methodology proposed by [16] was applied to obtain the Fitzpatrick Skin Type (FST) for each sample.

The FST was determined using the Individual Typology Angle (ITA) method. The ITA method provides a subjective-free approach for skin tone categorization, establishing a more objective and accurate criterion [17]. The method utilizes the  $L^*$  (Luminance) and  $b^*$  (yellow/blue) components from the CIELab color space, which is a perceptually uniform, device-independent color space defined by the International Commission on Illumination (CIE). The ITA value is calculated as the average of the pixel values within the analyzed area, as defined in Equation 1.

$$ITA = \arctan\left(\frac{L^* - 50}{b^*}\right) \times \frac{180^\circ}{\pi} \quad (1)$$

One of the challenges in determining ITA values is accurately capturing parts of the image that represent only the skin, without including areas containing any lesions. The lesion area, including dark edges (caused by the image capture equipment) and other unwanted artifacts, must be excluded from the calculation. Furthermore, the images are highly sensitive to factors such as lighting, quality, and other environmental variables.

In this work, the ITA method was applied to extract skin tone information from dermoscopic images, following the techniques outlined by [16]. Among the evaluated strategies, we adopt the structured cropping approach for patch selection, as it consistently yields superior results compared to central or random sampling. This method involves extracting samples from fixed peripheral regions of the image, deliberately avoiding the lesion area. As a result, it minimizes interference from lesion pixels in ITA calculations and enhances the representativeness of surrounding healthy skin. According to authors, this technique achieves lower mean absolute error (MAE) in ITA estimation and higher accuracy in Fitzpatrick skin type classification, while also offering greater robustness and reproducibility across images. Then, the ITA value for each image was calculated as the median of the values obtained (Equation 1).

## 4. Experiments

### 4.1 Scenarios

The experiments were carried out with careful attention to detail, aiming to identify the most effective combinations of extractors and classifiers across various scenarios, including different data augmentation techniques.

To ensure fair comparisons, identical configurations were applied to all scenarios, and the same test sets were used to evaluate model performance.

This was accomplished by stratifying the datasets, allocating 80% for training and 20% for testing. The results presented in each experiment represent the average of 5 runs, highlighting the thoroughness of the experimental process.

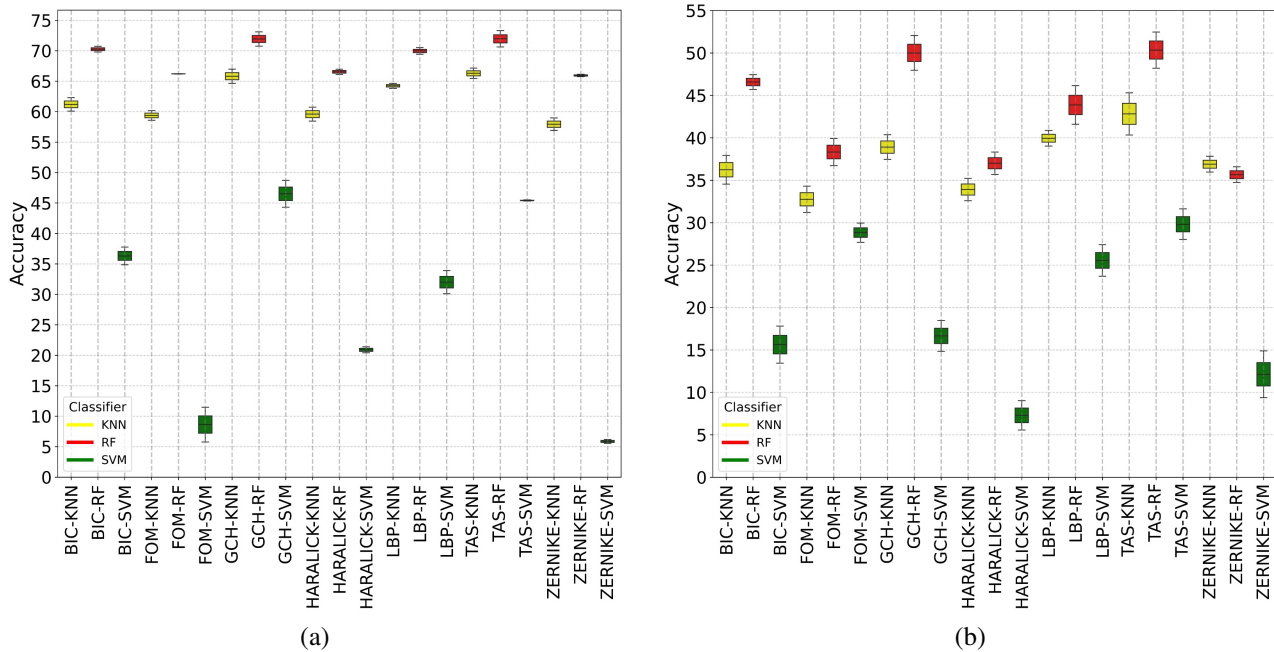


**Figure 2.** Mean accuracies  $\pm$  standard deviation obtained for each extractor-classifier combination for the datasets: (a) HAM10000 and (b) PAD-UFES-20, considering an imbalanced scenario without data augmentation. The results are visualized using boxplots to show the distribution of accuracies across runs.

## 4.2 Results and Discussion

Figures 2(a) and 2(b) show the accuracy results for each extractor-classifier combination on the HAM10000 and PAD-UFES-20 datasets, respectively, in a scenario without data augmentation and with significant class imbalance. We can notice that the SVM classifier performed poorly compared to the other classifiers ( $k$ -NN and RF) on both datasets. Overall, the RF classifier achieved the highest accuracies.

For the HAM10000 dataset (Figure 2(a)), the best extractor-classifier combinations were TAS-RF and GCH-RF, achieving accuracies of 71.95% and 71.83%, respectively. While these results were nearly identical, it is important to note that the TAS extractor generates 162 texture-based features, whereas the GCH extractor captures 30 color-based features.

For the PAD-UFES-20 dataset (Figure 2(b)), the top-performing combinations were also TAS-RF and GCH-RF, with average accuracies of 50.74% and 50%, respectively. Comparing the two datasets reveals a significant performance drop with the RF classifier for the PAD-UFES-20 dataset.

In addition to overall accuracies, analyzing the performance for each lesion class is essential. Figure 3 presents the confusion matrix for the best extractor-classifier combination (GCH-RF) applied to the HAM10000 and PAD-UFES-20 datasets, respectively. It is clear that the highest misclassification rates occur within the *nv* lesion class in the HAM10000 dataset, as well as the *bcc* and *akiec* lesion classes in the PAD-UFES-20 dataset.

Based on the performance analysis of each extractor and classifier, in an unbalanced scenario without data augmentation, some data augmentation techniques are explored. To

better evaluate the performance of each technique, the imbalance of the classes was maintained. Scenario A contained only original images, while scenarios B to H included only images generated by specific techniques. Additionally, scenarios H and I combined all techniques, with and without the inclusion of original images.

As shown in Figure 4, despite the overall increase in the average accuracy in scenarios H and I, there was no significant difference between the techniques. The GCH and TAS extractors consistently achieved the best performance across all scenarios.

After evaluating the impact of specific data augmentation techniques, initial analyses of skin tones were conducted using the original datasets (without augmentation). The ITA calculation, as detailed in Subsection 3.4, was used to categorize skin tones, enabling the assessment of sample distribution and potential bias related to skin tone.

Figures 5(a) and 5(b) illustrate the distribution of images in each of the 6 skin tones obtained by the calculation of ITA and structured region selection technique, from the datasets HAM10000 and PAD-UFES-20, respectively. The HAM10000 dataset (Figure 5(a)) shows a strong bias, with most images concentrated in Skin Tone Type 1 and underrepresentation of other tones. In contrast, the PAD-UFES-20 dataset (Figure 5(b)) has a more balanced distribution across skin tones, demonstrating better representativity.

Figures 6(a) and 6(b) further highlight the imbalance by showing the distribution of skin tones across different lesion classes in both datasets. The discrepancy in sample representation across tones and lesion types underscores the dataset

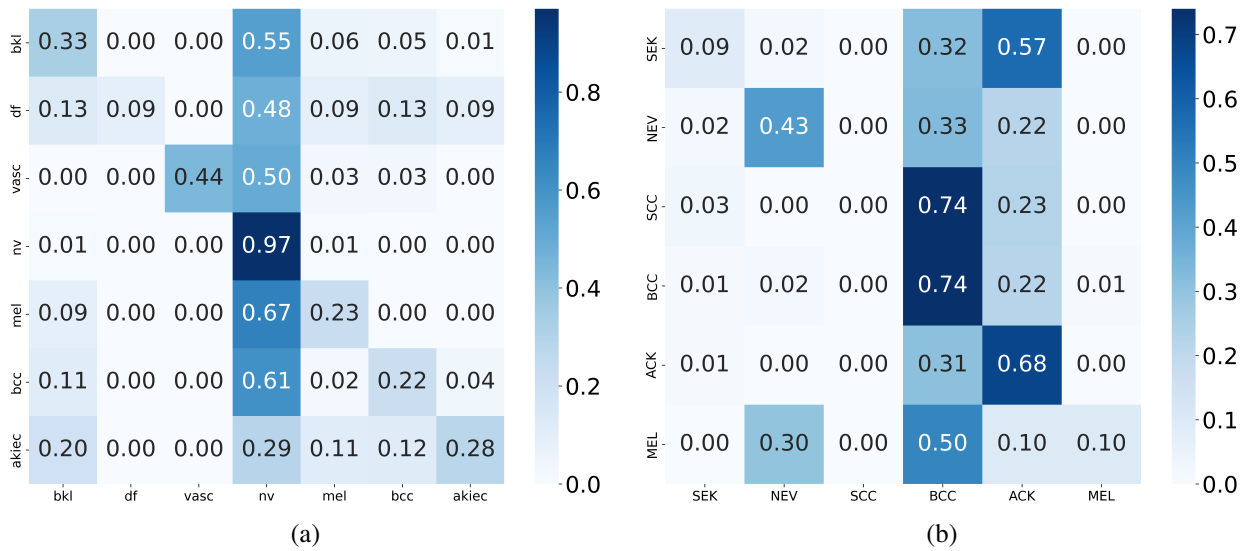


**Figure 3.** Confusion matrices (accuracies) obtained from the best extractor-classifier combination (GCH-RF) for each dataset: (a) HAM10000 and (b) PAD-UFES, considering an imbalanced scenario without data augmentation.

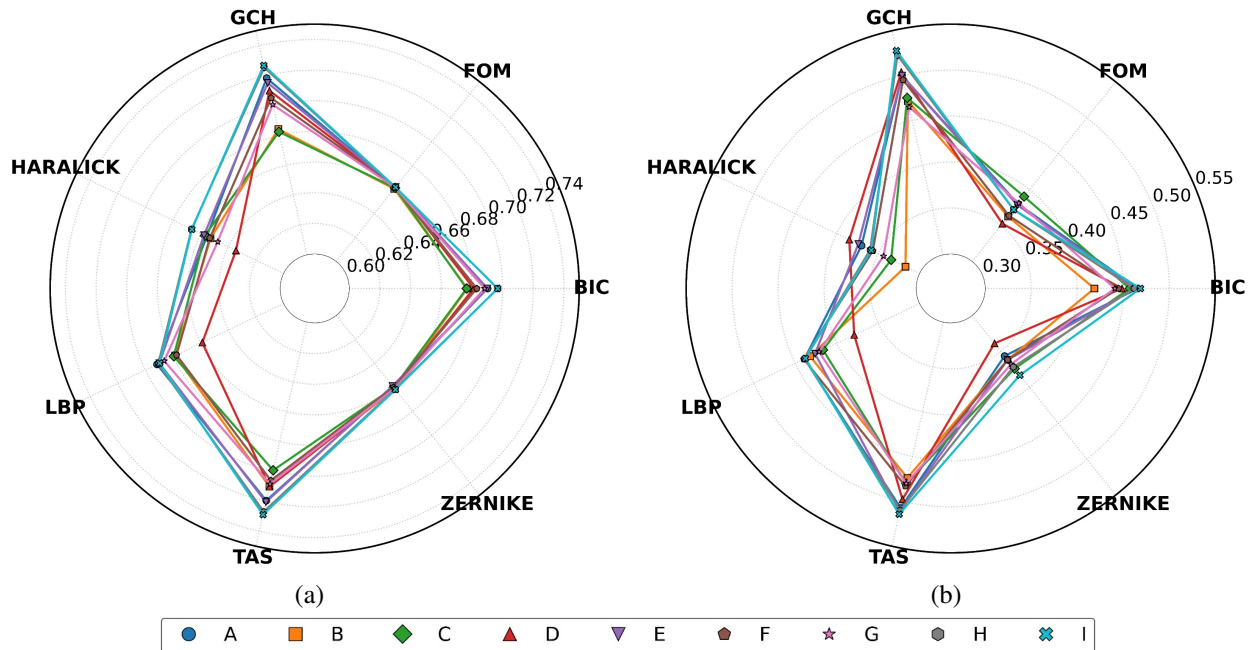


**Figure 4.** Performances obtained by each extractor with the RF classifier for the datasets: (a) HAM10000 and (b) PAD-UFES-20, considering imbalanced scenarios with data augmentation (scenarios A-I).

bias, which could impact classification performance.

For analysis of bias in relation to the sample distributions across each lesion and skin tone class, the proportions of errors and correct classifications obtained by the best extractor-classifier combination (GCH-RF) are presented, considering scenario I of data augmentation from the HAM10000 and PAD-UFES-20 datasets (Figures 7(a) and 7(b), respectively). For both datasets, the results highlight the impact of sample distribution across lesion and skin tone classes.

To ensure methodological rigor and prevent data leakage, we implemented strict patient- and lesion-level separation

protocols. All train-test splits were stratified by unique lesion identifiers, ensuring that images from the same lesion were never included in multiple subsets. Fixed global random seeds were used across all stochastic processes, and all classifiers were fully parameterized. For instance, the SVM utilized a defined search space (RBF/linear kernel;  $C$  [1, 10, 100, 1000]; gamma [1e-3, 1e-4]). Model selection was finalized based on cross-validation performance.

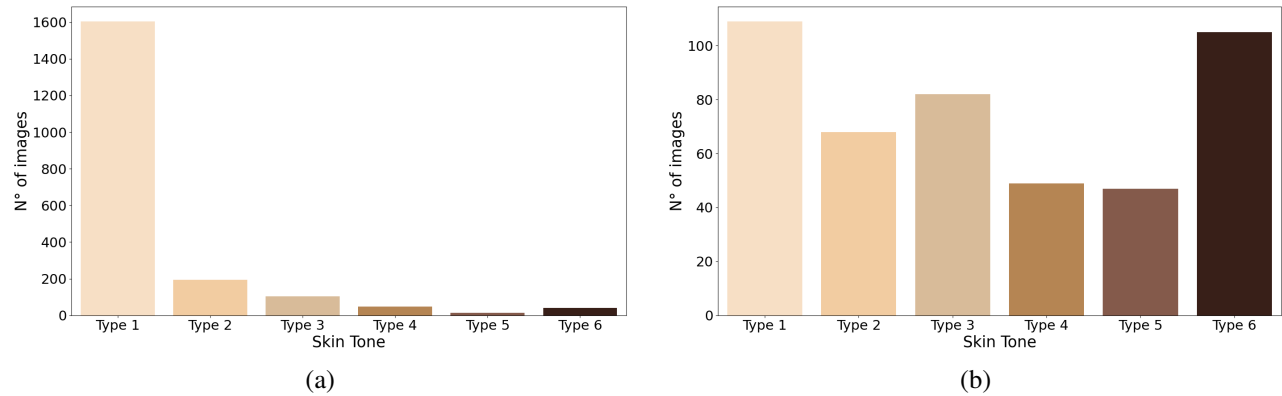


**Figure 5.** Distribution of images, in each of the 6 skin tones, obtained by calculating ITA and structured region selection technique, from the datasets: (a) HAM10000 and (b) PAD-UFES-20, respectively.

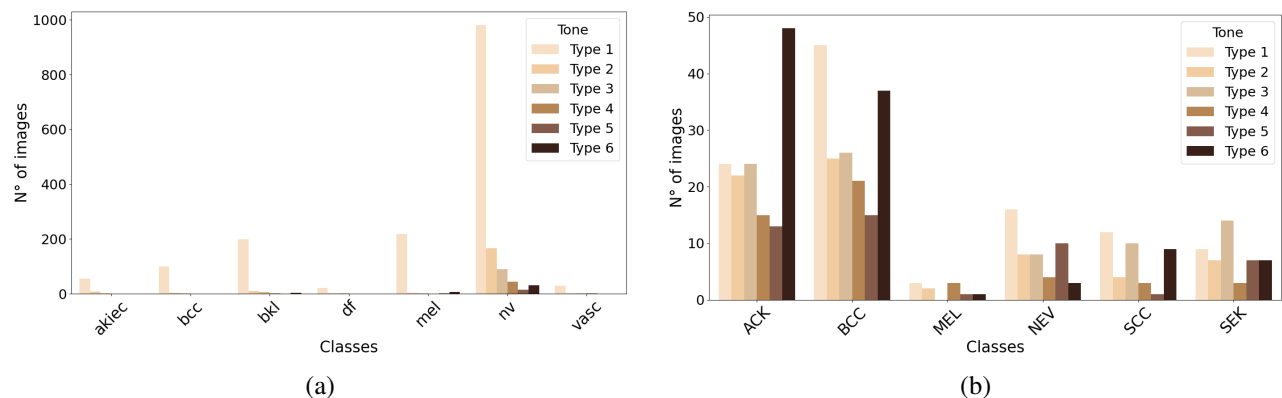


**Figure 6.** Distribution of images of (6) skin tones in each lesion class, obtained by calculating ITA and structured region selection technique, from the datasets (a) HAM10000 and (b) PAD-UFES-20, respectively.

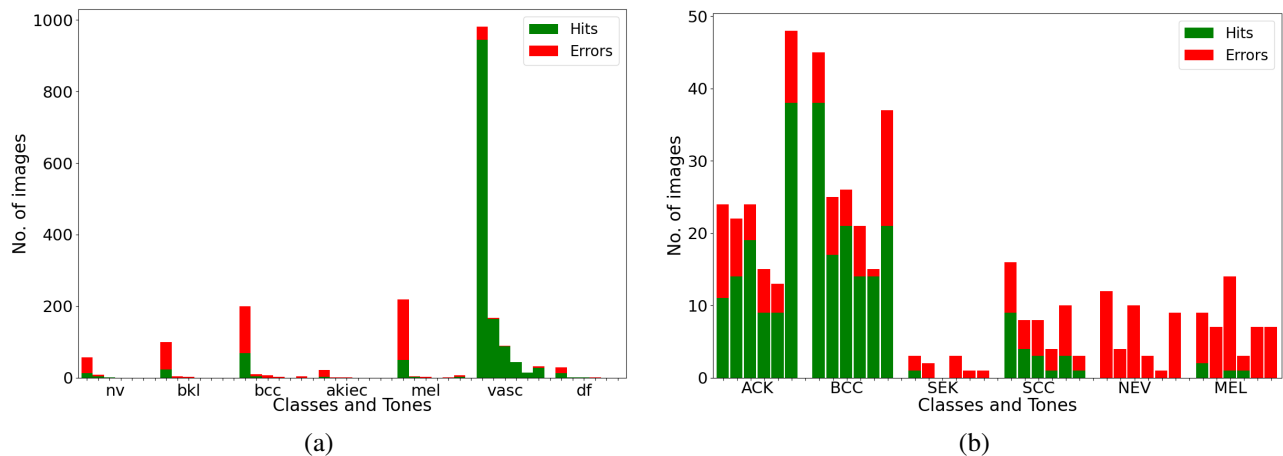


**Figure 7.** Distribution of the proportion of errors and correct classifications in each lesion and skin tone class, obtained using the best extractor-classifier combination GCH-RF, considering scenario I of data augmentation from the datasets: (a) HAM10000 and (b) PAD-UFES-20, respectively.

## 5. Conclusion

This work explores data augmentation scenarios in conjunction with bias analysis based on the distribution of skin tones and lesions, in order to contribute to discussions on skin lesion classification, particularly in contexts where certain skin tones

are underrepresented. The experiments also demonstrate that the best combination between feature extractor and classifier was the Global Color Histogram (GCH) with the Random Forest (RF) classifier.

For future work, we plan to investigate the use of active

learning and its impact on selecting more representative samples (in terms of skin tones and lesions) from the analyzed datasets.

## Acknowledgments

This work has been supported by CNPq; CAPES; FAPESP and UFSCar.

## Author Contributions

Thiago Meneses Lopes: Methodology, Code Development, Experiments, Analysis of Results, Writing (Original Draft and Review & Editing); Martin Heckmann: Methodology, Analysis of Results, Reviewed and Approved the Final Manuscript; Marcelo Ponciano da Silva: Assisted in Conceptualization, Methodology, Reviewed and Approved the Final Manuscript; Pedro Henrique Bugatti: Conceptualization, Methodology, Reviewed and Approved the Final Manuscript; Priscila Tiemi Maeda Saito: Conceptualization, Methodology, Analysis of Results, Writing (Review & Editing), Supervision.

## References

- [1] CHEN, J. Y. et al. Skin cancer diagnosis by lesion, physician, and examination type: a systematic review and meta-analysis. *JAMA Dermatology*, 2025.
- [2] SALINAS, M. P. et al. A systematic review and meta-analysis of artificial intelligence versus clinicians for skin cancer diagnosis. *NPJ Digital Medicine*, v. 7, n. 1, p. 125, 2024.
- [3] ALDRIDGE, R. B.; MAXWELL, S. S.; REES, J. L. Dermatology undergraduate skin cancer training: a disconnect between recommendations, clinical exposure and competence. *BMC Medical Education*, v. 12, n. 1, p. 27, 2012.
- [4] KINYANJUI, N. M. et al. *Estimating skin tone and effects on classification performance in dermatology datasets*. 2019. ArXiv preprint arXiv:1910.13268.
- [5] PEREZ, F. et al. Data augmentation for skin lesion analysis. In: *International Workshop on Computer-Assisted and Robotic Endoscopy (CARE)*. Cham: Springer International Publishing, 2018. p. 303–311.
- [6] RASHID, H.; TANVEER, M. A.; KHAN, H. A. Skin lesion classification using GAN-based data augmentation. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Berlin, Germany: IEEE, 2019. p. 916–919.
- [7] PHAM, T.-C. et al. Deep cnn and data augmentation for skin lesion classification. In: *Asian Conference on Intelligent Information and Database Systems (ACIIDS)*. Cham: Springer International Publishing, 2018. p. 573–582.
- [8] BOZKURT, F. Skin lesion classification on dermatoscopic images using effective data augmentation and pre-trained deep learning approach. *Multimed. Tools Appl.*, v. 82, n. 12, p. 18985–19003, 2023.
- [9] GALDRAN, A. et al. *Data-driven color augmentation techniques for deep skin image analysis*. 2017. ArXiv preprint arXiv:1703.03702.
- [10] KINYANJUI, N. M. et al. Fairness of classifiers across skin tones in dermatology. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Cham: Springer International Publishing, 2020. p. 320–329.
- [11] DANESHJOU, R. et al. Disparities in dermatology ai performance on a diverse, curated clinical image set. *Science Advances*, v. 8, n. 31, p. eabq6147, 2022.
- [12] BARROS, L.; CHAVES, L.; AVILA, S. Assessing the generalizability of deep neural networks-based models for black skin lesions. In: *Iberoamerican Congress on Pattern Recognition (CIARP)*. Cham: Springer Nature Switzerland, 2023. p. 1–14.
- [13] CORREA-MEDERO, R. L.; PATEL, B.; BANERJEE, I. Adversarial debiasing techniques towards ‘fair’ skin lesion classification. In: *2023 11th International IEEE/EMBS Conference on Neural Engineering (NER)*. Florence, Italy: IEEE, 2023. p. 1–4.
- [14] TSCHANDL, P.; ROSENDAHL, C.; KITTLER, H. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, v. 5, n. 1, p. 1–9, 2018.
- [15] PACHECO, A. G. C. et al. Pad-ufes-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones. *Data in Brief*, v. 32, p. 106221, 2020.
- [16] CORBIN, A.; MARQUES, O. Exploring strategies to generate fitzpatrick skin type metadata for dermoscopic images using individual typology angle techniques. *Multimedia Tools and Applications*, v. 82, n. 15, p. 23771–23795, 2023.
- [17] CHARDON, A.; CRETOIS, I.; HOURSEAU, C. Skin colour typology and suntanning pathways. *International Journal of Cosmetic Science*, v. 13, n. 4, p. 191–208, 1991.