Expert Agreement on the Presence and Spatial Localization of Melanocytic Features in Dermoscopy

Konstantinos Liopyris[1,2], Cristian Navarrete-Dechent[1,3], Michael A. Marchetti[1], Veronica Rotemberg[1], Zoe Apalla[4], Giuseppe Argenziano[5], Andreas Blum[6], Ralph P. Braun[7], Cristina Carrera[8,9], Noel C.F. Codella[10], Marc Combalia[8,9], Stephen W. Dusza[1], David A. Gutman[11,12], Brian Helba[13], Rainer Hofmann-Wellenhof[14], Natalia Jaimes[15,16], Harald Kittler[17], Kivanc Kose[1], Aimilios Lallas[4], Caterina Longo[18,19], Josep Malvehy[8,9], Scott Menzies[20,21], Kelly C. Nelson[22], John Paoli[23], Susana Puig[8,9], Harold S. Rabinovitz[24], Ayelet Rishpon[25], Teresa Russo[5], Alon Scope[26,27], H. Peter Soyer[28], Jennifer A. Stein[29], Willhelm Stolz[30], Dimitrios Sgouros[2], Alexander J. Stratigos[2], David L. Swanson[31], Luc Thomas[32], Philipp Tschandl[17], Iris Zalaudek[33], Jochen Weber[1], Allan C. Halpern[1] and Ashfaq A. Marghoob[1]

Dermoscopy aids in melanoma detection; however, agreement on dermoscopic features, including those of high clinical relevance, remains poor. In this study, we attempted to evaluate agreement among experts on exemplar images not only for the presence of melanocytic-specific features but also for spatial localization. This was a cross-sectional, multicenter, observational study. Dermoscopy images exhibiting at least 1 of 31 melanocytic-specific features were submitted by 25 world experts as exemplars. Using a web-based platform that allows for image markup of specific contrast-defined regions (superpixels), 20 expert readers annotated 248 dermoscopic images in collections of 62 images. Each collection was reviewed by five independent readers. A total of 4,507 feature observations were performed. Good-to-excellent agreement was found for 14 of 31 features (45.2%), with eight achieving excellent agreement (Gwet's AC >0.75) and seven of them being melanoma-specific features. These features were peppering/granularity (0.91), shiny white streaks (0.89), typical pigment network (0.83), blotch irregular (0.82), negative network (0.81), irregular globules (0.78), dotted vessels (0.77), and blue–whitish veil (0.76). By utilizing an exemplar dataset, a good-to-excellent agreement was found for 14 features that have previously been shown useful in discriminating nevi from melanoma. All images are public (www.isic-archive.com) and can be used for education, scientific communication, and machine learning experiments.

[1]*Dermatology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, USA;* [2]*Department of Dermatology, Andreas Syggros Hospital of Cutaneous & Venereal Diseases, University of Athens, Athens, Greece;* [3]*Department of Dermatology, Facultad de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile;* [4]*First Department of Dermatology, Aristotle University School of Medicine, Thessaloniki, Greece;* [5]*Dermatology Unit, University of Campania Luigi Vanvitelli, Naples, Italy;* [6]*Public, Private, and Teaching Practice of Dermatology, Konstanz, Germany;* [7]*Department of Dermatology, University Hospital Zürich, Zürich, Switzerland;* [8]*Melanoma Unit, Department of Dermatology, Hospital Clínic de Barcelona, University of Barcelona, Barcelona, Spain;* [9]*Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Valencia, Spain;* [10]*IBM Research AI, Thomas J. Watson Research Center, Yorktown Heights, New York, USA;* [11]*Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, Georgia, USA;* [12]*Department of Neurology, Emory University School of Medicine, Atlanta, Georgia, USA;* [13]*Kitware, Clifton Park, New York, USA;* [14]*Department of Dermatology, Medical University of Graz, Graz, Austria;* [15]*Dr. Phillip Frost Department of Dermatology and Cutaneous Surgery, Miller School of Medicine, University of Miami, Miami, Florida, USA;* [16]*Sylvester Comprehensive Cancer Center, University of Miami, Miami, Florida, USA;* [17]*Vienna Dermatologic Imaging Research Group, Department of Dermatology, Medical University of Vienna, Vienna, Austria;* [18]*Department of Dermatology, University of Modena and Reggio Emilia, Modena, Italy;* [19]*Centro Oncologico ad Alta Tecnologia Diagnostica, Azienda Unità Sanitaria Locale - IRCCS di Reggio Emilia, Reggio Emilia, Italy;* [20]*Faculty of Medicine and Health, Sydney Medical School, The University of Sydney, Camperdown, Australia;* [21]*Sydney Melanoma Diagnostic Centre, Royal Prince Alfred Hospital, Camperdown, Australia;* [22]*MD Anderson Cancer Center, Department of Dermatology, The University of Texas, Houston, Texas, USA;* [23]*Department of Dermatology and Venereology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden;* [24]*Department of Dermatology, University of Miami Miller School of Medicine, Miami, Florida, USA;* [25]*Department of Dermatology, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel;* [26]*Medical Screening Institute, Chaim Sheba Medical Center, Ramat Gan, Israel;* [27]*Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel;* [28]*Dermatology Research Centre, The University of Queensland Diamantina Institute, Brisbane, Australia;* [29]*The Ronald O. Perelman Department of Dermatology, New York University School of Medicine, New York, New York, USA;* [30]*Department of Dermatology, Ludwig-Maximilians-Universität, Munich, Germany;* [31]*Department of Dermatology, Mayo Clinic, Scottsdale, Arizona, USA;* [32]*Department of Dermatology, Centre Hospitalier de Lyon Sud, Hospices Civils de Lyon, Université Claude Bernard Lyon 1, Pierre Bénite, France; and* [33]*Dermatology Clinic, Maggiore Hospital, University of Trieste, Trieste, Italy*

Correspondence: Ashfaq A. Marghoob, Dermatology Service, Department of Medicine, Memorial Sloan Kettering Skin Cancer Center, 800 Veterans Memorial Highway, Hauppauge, New York 11788, USA. E-mail: marghooa@mskcc.org

Abbreviation: MLA, machine learning algorithm

Received 15 December 2022; accepted 19 January 2023; accepted manuscript published online 7 September 2023; corrected proof published online 11 January 2024

INTRODUCTION

Dermoscopy is a noninvasive, diagnostic technique that aids in the early diagnosis of melanoma (Achanta et al., 2010; Argenziano and Soyer, 2001; Argenziano et al., 2003). There is extensive literature highlighting the benefits of dermoscopy; however, although dermoscopy improves diagnostic accuracy, interobserver agreement for the presence of specific dermoscopic structures has been reported to be poor to moderate, even among experts (Achanta et al., 2012; Carrera et al., 2016). Reasons for subpar agreement are poorly characterized but could include differences in the perception of feature definitions among experts (Dinnes et al., 2018), potential overlap between features (Argenziano and Soyer, 2001), lack of standardization for dermoscopy terminology (Kittler et al., 2002), features that are intrinsically not reproducible (Argenziano et al., 2003), or use of inappropriate statistical measures of agreement (Carrera et al., 2016). Prior efforts in the field have attempted standardization of nomenclature; however, these terms have not been formally tested for reproducibility and reliability (Codella et al., 2018).

To achieve increased diagnostic accuracy and use among clinicians, agreement on dermoscopic structures and terminology standardization is necessary. To this objective, we designed the EASY (Expert Agreement Study on dermoscopY) of melanocytic lesions. In this study, we attempted to estimate the interobserver agreement of experts in dermoscopy on 31 melanocytic-specific dermoscopic features derived from the latest consensus agreement (Kittler et al., 2016). Toward this goal, dermoscopy experts were asked to submit exemplar lesions displaying any of the previously defined 31 melanocytic-specific structures (Kittler et al., 2016). These highly selective exemplar images were used to construct datasets to investigate the agreement of experts on the presence or absence (lesion-level agreement) of dermoscopic features. In addition, to our knowledge, a previously unreported superpixel (group of contextually similar pixels) annotation platform was created for readers to spatially annotate lesions by highlighting the superpixel containing specific structures. Whereas the lesion-level annotation allows for the analysis of interobserver agreement for the presence or absence of specific structures, the superpixel annotation platform permits analysis of interobserver spatial overlap and/or agreement.

RESULTS

Lesion-level annotations

Twenty experts annotated 248 images (eight images per exemplar feature), in groups of five annotators for a total of 4,507 feature markups. Single-reader observations of a feature accounted for 22.4% of all observations, whereas agreement of all readers evaluating an image for a feature occurred in 65 images (26.2%) (Supplementary Table S1).

Agreement on the exemplar feature on the lesion level. Measures of agreement are presented in Table 1. Percentage agreement varied by feature. The highest levels of percentage agreement were observed for peppering/granularity (92%), shiny white streaks (90%), typical network and irregular blotch (86%), negative network (84%), irregular globules and dotted vessels (82%), and scar-like depigmentation and blue–whitish veil (80%). The remaining 22 features yielded lower levels of percentage agreement (Table 1). Overall, Fleiss kappa showed poor agreement (<0.4) for all melanocytic-specific features with the exception of the rim of brown globules and irregular blotch (0.44 and 0.42, respectively). These results are due to the paradoxical nature of the statistic's performance at extremely low or high levels of feature prevalence. When using Gwet's AC, excellent agreement was observed for irregular globules (0.78), typical network (0.83), peppering/granularity (0.91), shiny white streaks (0.89), negative network (0.81), irregular blotch (0.82), blue–whitish veil (0.76), and dotted vessels (0.77). The remaining structures showed poor-to-moderate agreement.

After collapsing the individual features into supercategories on the basis of structural similarities, higher levels of overall percentage agreement and Gwet's AC were observed. Globules/clods, network, regression structures, shiny white structures, negative network, and vessels showed excellent agreement, with Gwet's AC values >0.81. Moderate agreement was observed for lines and structureless; dots and angulated lines yielded poor agreement.

Superpixel-level annotations

Each of the 248 images in our set was delineated into approximately 1,000 (mean = 1001.4, SD = 18.1) superpixel patches. A total of 47,524 superpixels were annotated by the expert readers in these images. Disagreement among readers occurred in 81.5% of the superpixels annotated (n = 38,732 superpixels)

Percentage superpixel agreement on the exemplar feature. There were nine features that yielded five readers' spatial agreement exceeding 10% of the superpixels where they were the exemplar feature (i.e., typical network with 36.2% absolute agreement among all five readers annotating the image [100%], cobblestone pattern with 24.7%, rim of brown globules with 19.6%, blotch regular with 17.6%, botch irregular with 17.2%, negative network with 15.3%, shiny white streaks with 15.2%, peppering/granularity with 11.3%, and blue–whitish veil with 11.3%) (Table 2 and Supplementary Table S4).

Ten features yielded no spatial agreement among the expert readers (0%); they included regular dots, milky red globules, regular globules, angulated lines, branched streaks, broadened network, delicate network, homogeneous pattern, milky red areas, and corkscrew vessels.

Percentage superpixel agreement on all features (exemplar and nonexemplar). The features with the highest level of absolute (100%) spatial agreement among readers were *cobblestone pattern* with 14.63% absolute (100%) agreement among readers, typical network with 11.88%, and rim of brown globules with 10.86% (Supplementary Table S2 and Supplementary Figure S3). The same 10 features that had no agreement in the exemplar group did not yield any spatial agreement among five readers when evaluating both exemplar and nonexemplar.

Confusion matrix and Dice coefficient (exemplar and nonexemplar features). Percentage agreement among experts on the superpixel level was comparatively low; however, there were 31 pairs of features displaying consistently high overlap, with an average Dice coefficient ≥0.5 (Table 3).

Table 1. Measures of Agreement for Individual Dermoscopic Features and Combined Superfeature Categories on a Lesion Level

Individual Dermoscopic Features				Combined Dermoscopic Superfeature Categories			
Variable	% Agreement	Fleiss' Kappa	Gwet's AC	Variable	% Agreement	Kappa	Gwet's AC
Dots:irregular	60.00%	0.0954	0.2829	Dots	61.00%	0.1371	0.2884[1]
Dots:regular	60.00%	0.2000	0.2000				
Globules/clods:cobblestone pattern	60.00%	0.0809	0.2918	Globules/clods	91.00%	0.3706[1]	0.8950[1]
Globules/clods:irregular	82.00%	−0.0989	0.7847[1]				
Globules/clods:regular	46.00%	−0.1315	−0.0301				
Globules/clods:rim of brown globules	78.00%	0.4419	0.6368[1]				
Lines:branched streaks	76.00%	0.0033	0.6839[1]	Lines	70.77%	0.3542[1]	0.4659[1]
Lines:pseudopods	54.00%	−0.1170	0.2180				
Lines:radial streaming	65.93%	0.0143	0.4792				
Lines:starburst (pseudopods/radial)	54.00%	0.0417	0.1154				
Network:atypical pigment network/reticulation	70.00%	−0.0631	0.5821[1]	Network	100%	1.0[1]	1.0[1]
Network:broadened pigment network/reticulation	38.00%	−0.2917	−0.1923				
Network:delicate pigment network/reticulation	48.00%	−0.0552	−0.0252				
Network:typical pigment network/reticulation	86.00%	0.2222	0.8293[1]				
Regression structures:peppering/granularity	92.00%	−0.0417	0.9133[1]	Regression structures	96.00%	-0.0204	0.9584[1]
Regression structures:scarlike depigmentation	80.00%	0.3225	0.7162[1]				
Shiny white structures:shiny white structures	90.00%	−0.0526	0.8895[1]	Shiny white structures	90.00%	-0.0526	0.8895[1]
Independent:angulated lines/polygons/zig-zag	60.00%	0.1883	0.2114	Angulated lines/polygons	60.00%	0.1883	0.2114
Independent:negative pigment network (independent)	84.00%	−0.0870	0.8124	Negative network	84.00%	-0.0870	0.8124
Structureless:blotch irregular	86.00%	0.4186	0.8156[1]	Structureless	75.06	0.1699[1]	0.6434[1]
Structureless:blotch regular	66.00%	0.2052	0.4058				
Structureless:blue-whitish veil	80.00%	−0.1111	0.7561[1]				
Structureless:milky red areas	58.00%	0.0349	0.2564				
Structureless:structureless brown (tan)	42.67%	−0.1485	−0.1448				
Structureless: Homogeneous:NOS	46.57%	−0.0876	−0.0876				
Vessels:comma	76.00%	0.0588	0.6779[1]	Vessels	89.49%	0.3462[1]	0.8748[1]
Vessels:corkscrew	58.00%	−0.1858	0.3496				
Vessels:dotted	82.00%	0.1477	0.7718[1]				
Vessels:linear irregular	68.89%	0.2742	0.4555[1]				
Vessels:polymorphous	54.00%	−0.0403	0.1753				
Vessels:milky red globules	67.00%	−0.0124	0.5104[1]				

Abbreviations: AC, agreement coefficient; NOS, not otherwise specified.
[1]Significant at $P < .05$

Atypical network and broadened network were annotated 85 times by different readers in superpixel regions with a Dice coefficient of 0.584, delicate network and typical network were annotated 67 times by different readers in superpixel regions with a Dice coefficient of 0.637, and broadened network and typical network were annotated 33 times by different readers in superpixel regions with a Dice coefficient of 0.658, showing the potential definition overlap among the four different features. Furthermore, homogeneous pattern—a global pattern—seemed to be nonspecific and overlapping with almost all other features annotated (Figure 1). Examples of both outcomes (high overlap/agreement in one feature but equally strong disagreement in other features) can be found in Supplementary Figure S1.

Lesion visualization. All images used in this study, along with their assorted annotations, have been made public and are available at http://multirater.isic-archive.com, whereas all images with their relevant metadata can be found at https://api.isic-archive.com/collections/166/.

DISCUSSION

In this study including 20 international dermoscopy experts, we assessed the agreement on 248 dermoscopic images for 31 established melanocytic-specific criteria (Kittler et al., 2016). Lesion-level agreement was moderate for many of the suggested features, with single-reader observations accounting for 22.4% of all feature selections. We found significant agreement for 14 of the 31 (45.2%) features examined, whereas excellent agreement was achieved on only 8 of the 31 features (25.8%), even when they were presented as exemplars. These features were peppering/granularity (0.91), shiny white streaks (0.89), typical pigment network (0.83), blotch irregular (0.82), negative network (0.81), irregular globules (0.78), dotted vessels (0.77), and blue–whitish veil (0.76) (Table 1). Interestingly, those with high agreement included 7 (50%) of the 14 melanoma-specific criteria examined, suggesting that they are pertinent to clinical practice (Yélamos et al., 2019). However, collapsing the features into supercategories of features on the basis of structural similarity yielded higher levels of

Table 2. Percentage Agreement among the Experts on the Spatial Localization (Superpixel Agreement) of the Exemplar Features

Feature	Mean Number of Readers	3RA %	5RA %
Dots:irregular	3.875	0.192648	0.003571
Dots:regular	2.5	0.250801	0
Globules/clods:cobblestone pattern	3.625	0.629329	0.247115
Globules/clodsIrregular	4.875	0.349432	0.043504
Globules/clods:milky red	1.25	0.016079	0
Globules/clods:regular	2.875	0.299281	0
Globules/clods:rim of brown globules	3.625	0.408223	0.195766
Lines:angulated lines/polygons/zig-zag pattern	3.5	0.156645	0
Lines:branched streaks	0.875	0	0
Lines:pseudopods	3.5	0.300592	0.003906
Lines:radial streaming	4	0.20898	0.035716
Network:atypical pigment network/reticulation	4.25	0.30728	0.031322
Network:broadened pigment network/reticulation	1.875	0.092118	0
Network:delicate pigment network/reticulation	2.125	0.01616	0
Network:negative pigment network	4.5	0.644936	0.152862
Network:typical pigment network/reticulation	4.375	0.756599	0.362415
Pattern:homogeneous: NOS	2.125	0.180564	0
Pattern:starburst	3.125	0.330246	0.041268
Regression structures:peppering/granularity	4.625	0.504966	0.112971
Regression structures:scarlike depigmentation	3.875	0.312567	0.050974
Shiny white structures:shiny white streaks	5	0.532319	0.151495
Structureless:blotch irregular	4.125	0.417691	0.172234
Structureless:blotch regular	3.375	0.508637	0.175758
Structureless:blue–whitish veil	4.375	0.4768	0.112656
Structureless:milky red areas	3	0.184145	0
Structureless:structureless brown (tan)	2.5	0.157379	0.033333
Vessels:comma	4.375	0.176428	0.027243
Vessels:corkscrew	1.125	0	0
Vessels:dotted	4.5	0.541738	0.082328
Vessels:linear irregular	3.875	0.247682	0.036885
Vessels:polymorphous	3.25	0.315864	0.002083

Abbreviations: 3RA %, 3 readers annotating the same region %; 5RA %, 5 readers annotating the same region %; NOS, not otherwise specified.

The mean number of readers shows the average n of readers annotating the images for this feature. The percentage of agreement ≥60% (at least 3 of 5 readers annotating the same region/superpixels for the same feature) can be seen under 3RA %, whereas under full 5RA %, we can see the percentage of absolute agreement for each feature (5 of 5 readers annotating the same region/superpixels for the same feature).

agreement, with the exemption of dots. These findings suggest that there could be an overlap among definitions and the perception of these features among experts that could be further elucidated to create more reliable diagnostic criteria. As expected, positive identification by the expert readers of the 31 melanocytic-specific features was higher when they were presented as exemplars than when nonexemplar features were identified in images by the experts (Supplementary Figure S2).

Prior efforts in the field have shown poor-to-moderate agreement for most dermoscopic features (Argenziano et al., 2003; Carrera et al., 2016; Kittler et al., 2016); however, what distinguishes our study is the use of exemplar images, submitted by the experts who initially described the majority of these features. The use of exemplar images yielded a better agreement than the prior efforts and has created a publicly available exemplar dataset (International Skin Imaging Collaboration). In addition, this study also utilizes spatial localization on the basis of the superpixel concept for specific dermoscopic features. This, to our knowledge, previously unreported approach allowed for the identification of overlapping features, which may offer important insights into the visual perception of features by readers.

Overall agreement and chance-corrected agreement were highly variable for specific dermoscopic features. However, an important aspect of our study was the use of exhaustive (superpixel) annotations for the feature localization on a dermoscopic image. Superpixel annotations allow for additional refinement of our understanding of dermoscopic features and alternative approaches to feature agreement analysis (Figure 2). Agreement on the superpixel level was low; only 19.6% of all superpixels annotated showed any agreement between at least two readers, and only 11 features achieved an agreement >2% on the superpixel level among five readers. However, in our study, we identified 50 pairs of features that displayed constantly a high (>0.5) Dice overlap, highlighting the features that may perceptually be confused with each other or have definitions that may overlap with each other, suggesting that differing terminology may in fact be referring to the same feature or different features that occur concomitantly (Table 3 and Figure 1). We included four different features termed network/reticulation with

Table 3. Pairs of Features with High (>0.5) Dice Spatial Overlap on the Superpixel Annotations

n of pairs	% Overlap	Feature 1	Feature 2
92	52.80%	Network:atypical pigment network/reticulation	Network:typical pigment network/reticulation
85	58.40%	Network:atypical pigment network/reticulation	Network:broadened pigment network/reticulation
68	55.60%	Vessels:linear irregular	Vessels:polymorphous
67	63.70%	Network:delicate pigment network/reticulation	Network:typical pigment network/reticulation
64	67.00%	Lines:radial streaming	Pattern:starburst
58	86.50%	Globules/clods:cobblestone pattern	Globules/clods:regular
42	90.10%	Pattern:homogeneous: NOS	Structureless:blotch regular
33	65.80%	Network:broadened pigment network/reticulation	Network:atypical pigment network/reticulation
31	71.50%	Globules/clods:irregular	Globules/clods:regular
29	66.30%	Pattern:homogeneous: NOS	Structureless:structureless brown (tan peripheral area)
29	50.70%	Network:negative pigment network/reticulation	Shiny white structures:shiny white streaks
25	62.20%	Dots:regular	Globules/clods:regular
24	67.90%	Globules/clods:regular	Globules/clods:rim of brown globules
24	56.20%	Pattern:homogeneous:NOS	Structureless:blotch irregular
22	50.30%	Dots:irregular	Dots:regular
18	70.10%	Lines:pseudopods	Pattern:starburst
16	54.90%	Pattern:homogeneous:NOS	Structureless:blue–whitish veil
14	72.30%	Structureless:blotch regular	Structureless:blue–whitish veil
14	67.00%	Globules/clods:rim of brown globules	Lines:pseudopods
13	50.30%	Structureless:blotch regular	Structureless:structureless brown (tan
12	57.50%	Dots:regular	Network:atypical pigment network/reticulation
9	56.80%	Globules/clods:milky red	Vessels:polymorphous
9	52.40%	Globules/clods:regular	Network:negative pigment network/reticulation
7	92.90%	Structureless:blotch irregular	Structureless:blotch regular
5	71.60%	Lines:branched streaks	Pattern:starburst
5	61.40%	Globules/clods:cobblestone pattern	Pattern:homogeneous:NOS
5	50.00%	Vessels:corkscrew	Vessels:polymorphous
4	89.40%	Dots:regular	Globules/clods:cobblestone pattern
4	74.10%	Globules/clods:cobblestone pattern	Network:negative pigment network
4	72.30%	Globules/clods:rim of brown globules	Pattern:starburst
3	50.30%	Regression structures:peppering/granularity	Structureless:blotch regular

Abbreviation: NOS, not otherwise specified.

n of pairs indicates how many times those pairs occurred with high overlap in our study, whereas % overlap shows the percentage of their spatial overlap.

structural similarities based on the latest dermoscopy terminology consensus (Kittler et al., 2016); two of them are melanoma-specific features, which have been correlated with the diagnosis of melanoma on dermoscopy (i.e., atypical pigment network or for melanoma 2.0–9.0 and broadened pigment network): one is correlated with the diagnosis of nevi on dermoscopy (i.e., typical pigment network) and one nonspecific (i.e., delicate pigment network) (Kittler et al., 2016; Yélamos et al., 2019). However, despite their different definitions and significance in clinical practice, these features showed consistently a high Dice overlap in our study. Furthermore, this was also the case for features globules/clods:irregular (melanoma-specific feature or for melanoma 1.7–4.8) and globules/clods:regular, which is a feature specific for benign melanocytic lesions (Table 3 and Figure 1) (Kittler et al., 2016; Yélamos et al., 2019). Future studies are needed to explore whether these findings can be attributed to a different perception of overlapping features owing to ambiguous definitions of these features, different cognitive perceptions of features, or frequent coexistence of these features.

The superpixel approach has been previously used for the training and testing of machine learning algorithms (MLA); however, the ground truth for this task was set by a single annotator, and the performance of MLA was poor (Codella et al., 2019[1], 2018). MLA have shown potential for diagnosing melanoma and other skin cancers on dermoscopic images and could potentially prove to be an adjunct to clinical practice (Marchetti et al., 2020; Tschandl et al., 2020, 2019). Our results and the lack of agreement between annotators for the localization of features suggest that a single annotation for dermoscopic features may not be sufficient to determine ground-truth annotations owing to concerns related to reproducibility. We demonstrated that multiple reader annotations of curated datasets of dermoscopic images can help improve our understanding of image features and inform and improve MLA performance through supervised or active learning.

The primary goal of our study was to explore expert agreement for the 31 melanocytic-specific dermoscopic features; our results show the features that are more likely to be agreed upon and which features may overlap with others.

[1] Codella N, Rotemberg V, Tschandl P, Celebi ME, Dusza S, Gutman D, et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the International Skin Imaging Collaboration (ISIC). arXiv 2019.

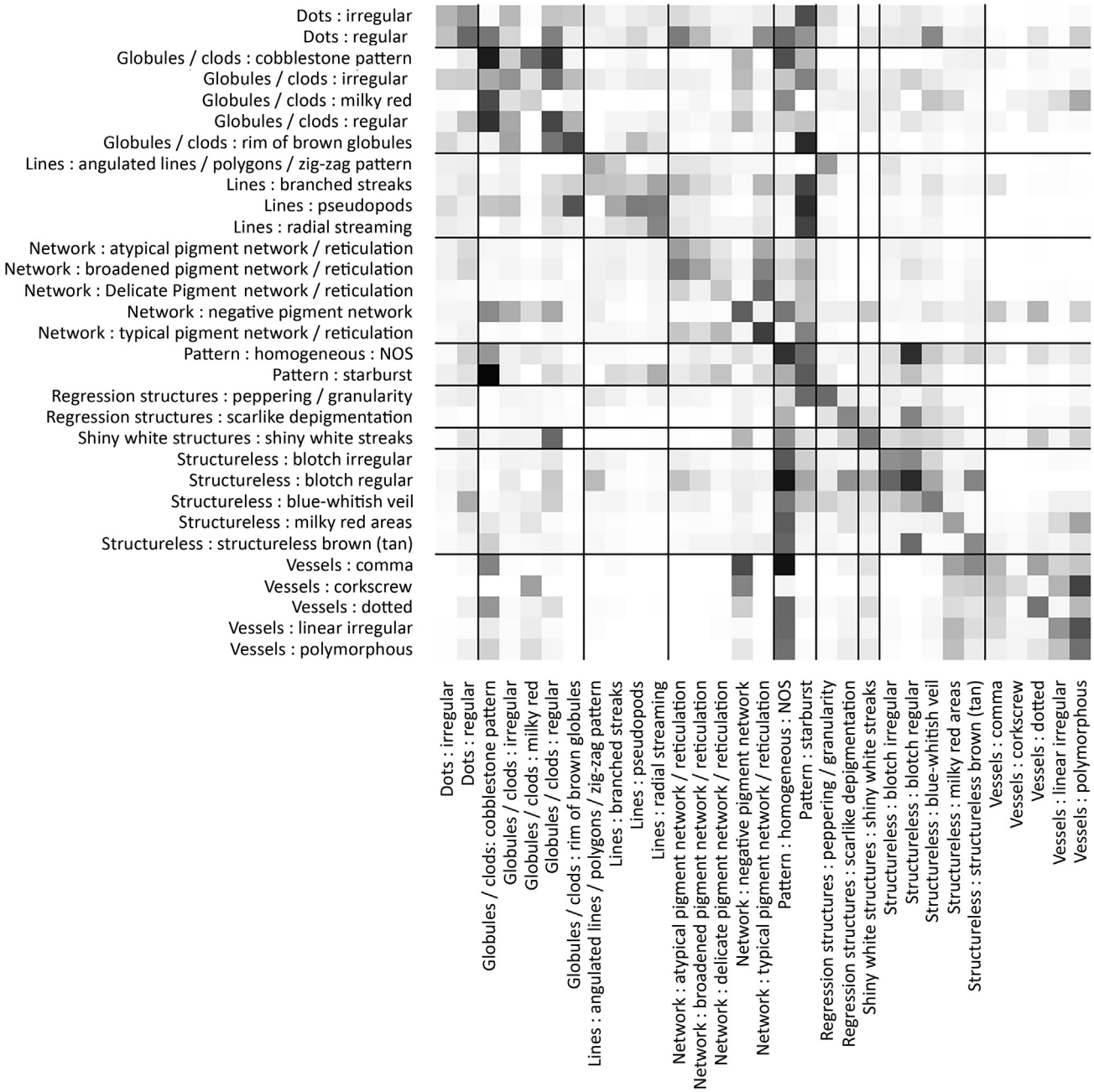


Figure 1. Confusability matrix. Each element of this confusion matrix is the median Dice coefficient across all pairs of features that occurred in the dataset. In the cases where the diagonal elements are as close to 1.0 as possible (dark color), this suggests that all pairs of readers (each individual reader directly compared with each of the other four, across all studies) selected close to identical sets of superpixels for the same feature and image. When all off-diagonal elements approach 0.0 (light color), this suggests that for different features, pairs of readers rarely selected overlapping superpixels. Feature pattern:homogeneous, a nonspecific feature that can be observed in any region of the lesion overlaps with a multitude of features, commonly with blotch:regular and vessels:comma; however, the agreement on the superpixel level for this feature among experts for this feature was 0%. In addition, features dots:regular and dots:irregular overlap among them and with globules/clods, whereas both the superpixel level agreement and Gwet's AC were low for these features.

This type of study can be used as the basis to evaluate which features are reliable and reproducible and can provide a critical basis for improved diagnostic algorithms by creating a standardized terminology for international consensus, potentially leading to higher diagnostic accuracy both by clinicians and by MLAs. An important goal of our study was to create a gold-standard dataset of exemplar images displaying the melanocytic-specific criteria for medical teaching, effective scientific communication, and machine learning experiments. There are now 248 dermoscopic images publicly available at http://multirater.isic-archive.com with their assorted annotations and areas of agreement for the 31 melanocytic-specific criteria, generated by experts in dermoscopy.

In conclusion, agreement on dermoscopic features remains variable, even when using highly selected exemplar images

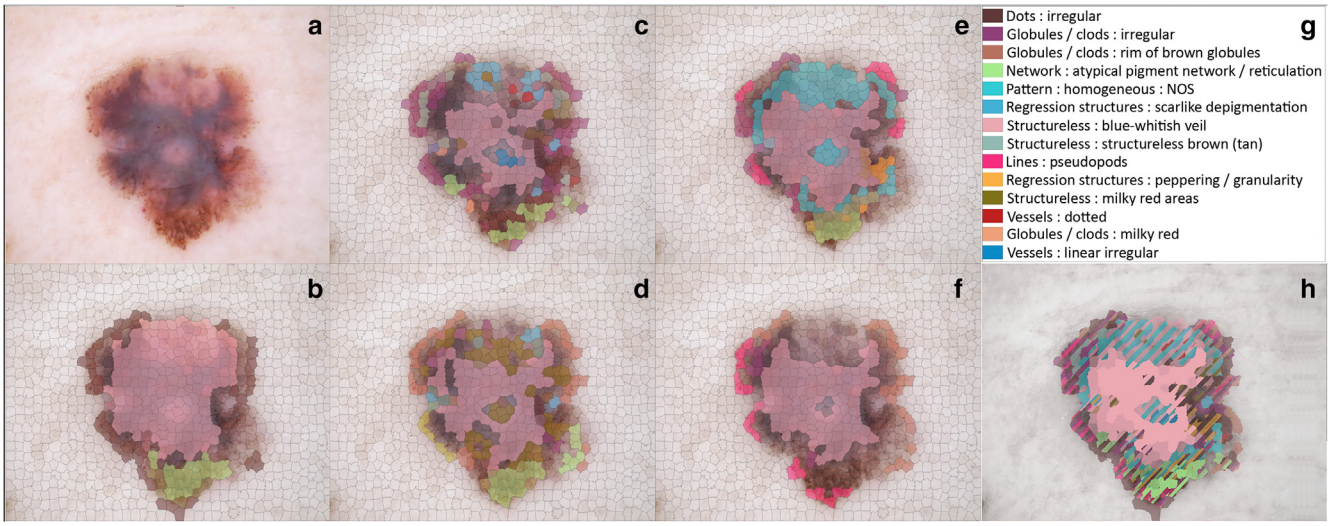


Figure 2. Superpixel annotation of exemplar image. (a) Sample image (ISIC_0022328) of a melanoma included in the study for exemplar feature structureless:blue–whitish veil. (**b**–**f**) Image overlaid with the superpixel outlines and annotation markups selected by each of the five expert readers. (**g**) Color representation of each feature selected by the readers. (**h**) Agreement overlaps among all readers.

among experts. The use of exemplar images facilitates the agreement process, whereas the use of superpixels provides insights into commonly overlapping features and the intrinsic intricacies of dermoscopic features. We found that half of the melanoma-specific criteria have good reproducibility, yielding high agreement, whereas others, such as broadened pigment network, delicate pigment network, and tan peripheral brown areas, have very low agreement even on the best case scenario of exemplar images and may need elimination from the lexicon of dermoscopy. Finally, analysis of the superpixel spatial overlap revealed that only 19.6% of superpixel overlaps occurred between different features, indicating that there may be redundancy or confusion in feature terminology. These observations are the necessary prerequisites for enhanced consensus building and nomenclature standardization. Our results will be used as a roadmap to achieve better terminology standardization and to guide melanoma diagnostic algorithms that ultimately would increase worldwide use and acceptance of dermoscopy.

Limitations

The use of superpixels is a different approach for the evaluation of agreement. However, these predetermined regions might not be optimal for attribute identification because dermoscopic features may not be bounded by superpixels or contrast thresholds, especially features such as networks, which inherently contain high contrast. Another important limitation of this study is the use of expert readers; extrapolating the results of this study to the general dermatology community must be made with caution. Another limitation of this study is that we did not balance the images for skin tone of the patients; however, including a diverse dataset from different parts of the world enhances the variety of skin tumor presentation. Finally, this should not be interpreted as an epidemiologic study because we have intentionally enriched the dataset for melanomas and their exemplar features.

MATERIALS AND METHODS

This cross-sectional, observational study was performed between September 1, 2017 and January 31, 2020. Invites were sent through email to 32 experts in dermoscopy. Experts were defined as clinicians with >10 years of experience in dermoscopy who have made significant contributions to research or teaching dermoscopy of pigmented lesions. We requested contributions of 1–3 exemplar-quality images for each of the 31 dermoscopic features (Kittler et al., 2016) (see Supplementary Table S3 for feature definitions and abbreviations). An exemplar-quality image was defined as one with excellent quality and an in-focus depiction of a feature in which the experts had absolute confidence regarding its presence in the lesion. In total, 25 experts (81%) contributed 964 images of melanocytic lesions for the 31 features. Of the 32 experts initially invited to participate in the study, 21 (66%) completed the annotation phase. Five additional experts were subsequently invited to participate in image annotation, and all agreed to participate. The images were uploaded to the International Skin Imaging Collaboration Archive and are available online (Supplementary Table S1 and http://multirater.isic-archive.com).

To determine the required number of readers and evaluations per reader, we used Monte Carlo simulations of the intraclass correlation coefficient. We estimated that with a confidence interval of 95% and an interclass correlation coefficient of 0.5 as the measure of agreement, five readers per study dataset would be efficient for evaluating agreement for the 31 dermoscopic features.

Feature selection

All 31 melanocytic-specific dermoscopic features from the 2016 International Dermoscopy Society terminology consensus were selected (Kittler et al., 2016). This list included 14 criteria previously reported to be significantly associated with melanoma diagnosis (Yélamos et al., 2019).

Dataset creation

Three experts (KL, CN-D, and AAM) selected 310 of the contributed images through consensus; 10 images per exemplar feature were selected on the basis of image quality and the presence of all

assorted metadata. These images were used to generate five splits each containing 62 images. Each split contained two exemplar images for each of the 31 dermoscopic features. The 25 expert readers were randomly allocated into groups of five and assigned to one of the five datasets on which to perform the annotation process. Image annotations were not completed for one of the splits owing to lack of participation; for the fifth dataset, only three of five experts completed annotations (23 of 25 experts completed annotations, 92%), and it was therefore excluded from further analysis.

Superpixel generation

After the images were uploaded to the International Skin Imaging Collaboration Archive (www.isic-archive.com), a superpixel map was automatically generated. Superpixel segmentation is a technique of grouping together the pixels of a given image into patches, according to their contextual similarity and spatial proximity. We applied SLIC (Simple Linear Iterative Clustering) algorithm on the pixel color values to create grouping distinctions on the basis of contrast thresholds (Achanta et al., 2012, 2010) (Figure 2).

Image annotations

We developed an annotation tool that allows for selection of features within images both on the lesion level (lesion-level annotations) and on the superpixel level (spatial annotations) (International Skin Imaging Collaboration). Lesion-level annotations allow for annotation of descriptive features in a binary fashion—present or not present—without indicating their exact location within the lesion. In contrast, superpixel-level annotation permits for spatial localization of features within the lesion. The annotation tool can be used to identify multiple overlapping features for any given lesion image. Annotators determined whether a feature was present or absent in a given image and in a given superpixel, according to their perception (video demonstration of annotations can be found at https://youtu.be/jgJdCD3k3Es).

Expert readers were blinded to the diagnosis and were invited to exhaustively annotate all 31 melanocytic-specific dermoscopic features in images. Readers were not instructed on the definitions of the features and annotated the features at their discretion. An example image displaying the superpixel division of an image and the assorted annotations performed by all five readers along with their agreement is displayed in Figure 2.

Agreement analysis

We performed three levels of analysis: (i) agreement for the exemplar feature on the lesion level, (ii) agreement for the exemplar features on the superpixel level, and (iii) agreement for the nonexemplar features on the superpixel level.

Lesion-level agreement

Agreement on the presence of exemplar feature. Exemplars of each of the 31 features were allocated to each of the four splits (two exemplar images per feature, per dataset). Data were combined into a single dataset for analysis. Percentage agreement, Fleiss kappa, and Gwent's AC1 were estimated for each of the 31 dermoscopic features (Wongpakaran et al., 2013). Both the Fleiss kappa and Gwet's AC1 were estimated owing to the paradoxical performance of the Fleiss kappa at the extremes of the distribution of percentage agreement. To further explore potential overlap among terms regarding similar features, these were combined into seven supercategories of features (i.e., dots, globules/clods, lines, network, regression structures, structureless, and vessels), maintaining separately the structurally distinct features of shiny white structures, angulated lines/polygons, and negative network on the basis of consensus among four investigators (KL, CN-D, MAM, and AAM). Measures of agreement were estimated for the supercategories. Kappa and Gwet's AC1 were interpreted as outlined by Landis and Koch: $0 < 0.4$ (poor agreement), $0.4 < 0.75$ (fair to good), and 0.75−1.0 (excellent agreement) (Landis and Koch, 1977).

Superpixel-level agreement

To assess inter-reader agreement on the same superpixel as well as the confusability or overlap with other features, we computed both the percentage agreement for each feature as well as the Dice coefficient for each possible cross-rater pair of selected superpixel sets (i.e., for both the same feature as well as different features marked by two readers) (Dice, 1945; Zijdenbos et al., 1994). For the percentage agreement, we took the total number of superpixels annotated within an image for a specific feature as the denominator (100%), and subsequently, we calculated the agreement for each feature on the annotated images. For the Dice coefficient, our calculations yielded a number between a minimum of 0.0 (0%) in cases of full disagreement (i.e., two mutually exclusive sets of superpixels) and a maximum of 1.0 (100%) in cases of complete agreement (i.e., exactly the same set of superpixels). To visually represent the full matrix of feature pair annotations for all study participants, we created a confusion matrix. Code in the Python (version 3.7) language for these analyses is available at Github repository (International Skin Imaging Collaboration Archive access Python module).

DATA AVAILABILITY STATEMENT

All data for this study have been made publicly available at http://multirater.isic-archive.com, whereas images and their assorted metadata can be found at https://api.isic-archive.com/collections/166/, and the Python code for the analyses performed can be found at https://github.com/neuroelf/isicarchive/blob/master/EASY - MARk-SUP - analyses.ipynb.

ORCIDs

Konstantinos Liopyris: http://orcid.org/0000-0001-9566-8238
Cristian Navarrete-Dechent: http://orcid.org/0000-0003-4040-3640
Michael A. Marchetti: http://orcid.org/0000-0002-1793-1851
Veronica Rotemberg: http://orcid.org/0000-0003-0639-2677
Zoe Apalla: http://orcid.org/0000-0002-9255-8196
Giuseppe Argenziano: http://orcid.org/0000-0003-1413-8214
Andreas Blum: http://orcid.org/0000-0003-4087-5015
Ralph P. Braun: http://orcid.org/0000-0003-1253-8206
Cristina Carrera: http://orcid.org/0000-0003-1608-8820
Noel C. F. Codella: http://orcid.org/0000-0001-6735-9067
Marc Combalia: http://orcid.org/0000-0001-5237-4256
Stephen W. Dusza: http://orcid.org/0000-0002-0747-2479
David A. Gutman: http://orcid.org/0000-0002-0792-7664
Brian Helba: http://orcid.org/0000-0003-2628-805X
Rainer Hofmann-Wellenhof: http://orcid.org/0000-0001-8688-3243
Harald Kittler: http://orcid.org/0000-0002-0051-8016
Kivanc Kose: http://orcid.org/0000-0003-3185-2639
Aimilios Lallas: http://orcid.org/0000-0002-7193-0964
Caterina Longo: http://orcid.org/0000-0002-8218-3896
Josep Malvehy: http://orcid.org/0000-0002-6998-914X
Scott Menzies: http://orcid.org/0000-0001-6430-6429
Kelly C. Nelson: http://orcid.org/0000-0002-5042-7366
John Paoli: http://orcid.org/0000-0003-1326-8535
Susana Puig: http://orcid.org/0000-0003-1337-9745
Harold S. Rabinovitz: http://orcid.org/0000-0001-6895-9697
Ayelet Rishpon: http://orcid.org/0000-0003-0555-4822
Teresa Russo: http://orcid.org/0000-0003-1792-9716
Alon Scope: http://orcid.org/0000-0001-9160-7411
H. Peter Soyer: http://orcid.org/0000-0002-4770-561X
Jennifer A. Stein: http://orcid.org/0000-0002-0962-4982
Dimitrios Sgouros: http://orcid.org/0000-0003-0009-8821
Alexander J. Stratigos: http://orcid.org/0000-0002-6088-5694
David L. Swanson: http://orcid.org/0000-0002-6682-7364

Luc Thomas: http://orcid.org/0000-0002-4315-5374
Philipp Tschandl: http://orcid.org/0000-0003-0391-7810
Iris Zalaudek: http://orcid.org/0000-0002-7878-4955
Jochen Weber: http://orcid.org/0000-0002-2397-4954
Allan C. Halpern: http://orcid.org/0000-0001-7320-1901
Ashfaq A. Marghoob: http://orcid.org/0000-0001-6068-0114

CONFLICT OF INTEREST

VR is an expert consultant to Inhabit Brands. NCFC is currently an employee of Microsoft and previously an employee of IBM during the conduct of a portion of the work. He has diverse investments in the healthcare and technology sectors. HK received speaker fees from Fotofinder, Eli Lilly, and Novartis and has equipment with Heine, Fotofinder, 3-Gen, and Casio. JM is a cofounder of Athena Tech s.l. HSR is a consultant for 3-Gen, Metaoptima, Dermasensor, Casio, and Caliber ID. HPS is a shareholder of MoleMap NZ and e-derm Consult GmbH and undertakes regular teledermatological reporting for both companies. He is a medical consultant for Canfield Scientific, MoleMap Australia Pty, and Blaze Bioscience and a medical advisor for First Derm. AAM received an honorarium for lecturing on the subject of dermoscopy from DermLite, Canfield, Heine, and FotoFinder. The remaining authors state no conflict of interest.

ACKNOWLEDGMENTS

The group would like to acknowledge the International Dermoscopy Society for its efforts and aid in the completion of this study. This work was supported in part through the Memorial Sloan Kettering Cancer Center institutional National Institutes of Health/National Cancer Institute Cancer Center Support Grant P30 CA008748.

AUTHOR CONTRIBUTIONS

Conceptualization: KL, CN-D, MAM, JW, SWD, VR, ACH, AAM; Data Curation: KL, CN-D, MAM, JW, ACH, AAM, VR, ZA, HA, AB, RPB, CC, MC, SWD, DAG, BH, RH-W, NJ, HK, KK, AL, CL, HM, SM, KCN, JP, SP, HSR, AR, TR, AS, HPS, JAS, WS, DS, AJS, DLS, LT, PT, IZ, JW

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2023.01.045.

REFERENCES

Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk A. SLIC superpixels. https://infoscience.epfl.ch/record/149300?ln=en; 2010. (accessed November 22, 2023).

Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans Pattern Anal Mach Intell 2012;34:2274–82.

Argenziano G, Soyer HP. Dermoscopy of pigmented skin lesions–a valuable tool for early diagnosis of melanoma. Lancet Oncol 2001;2:443–9.

Argenziano G, Soyer HP, Chimenti S, Talamini R, Corona R, Sera F, et al. Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet. J Am Acad Dermatol 2003;48:679–93.

Carrera C, Marchetti MA, Dusza SW, Argenziano G, Braun RP, Halpern AC, et al. Validity and reliability of dermoscopic criteria used to differentiate nevi from melanoma: a web-based International Dermoscopy Society study. JAMA Dermatol 2016;152:798–806.

Codella NCF, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW, 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). A paper presented at: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). 4–7 April 2018, Washington, DC.

Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26:297–302.

Dinnes J, Deeks JJ, Chuchu N, Ferrante di Ruffano L, Matin RN, Thomson DR, et al. Dermoscopy, with and without visual inspection, for diagnosing melanoma in adults. Cochrane Database Syst Rev 2018;12:CD011902.

International Skin Imaging Collaboration (ISIC). Archive. www.isic-archive.com. (accessed December 4, 2023).

Kittler H, Marghoob AA, Argenziano G, Carrera C, Curiel-Lewandrowski C, Hofmann-Wellenhof R, et al. Standardization of terminology in dermoscopy/dermatoscopy: results of the third consensus conference of the International Society of dermoscopy. J Am Acad Dermatol 2016;74:1093–106.

Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. Lancet Oncol 2002;3:159–65.

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

Marchetti MA, Liopyris K, Dusza SW, Codella NCF, Gutman DA, Helba B, et al. Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma: results of International Skin Imaging Collaboration 2017. J Am Acad Dermatol 2020;82:622–7.

Tschandl P, Codella N, Akay BN, Argenziano G, Braun RP, Cabo H, et al. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. Lancet Oncol 2019;20:938–47.

Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human–computer collaboration for skin cancer recognition. Nat Med 2020;26:1229–34.

Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 2013;13:61.

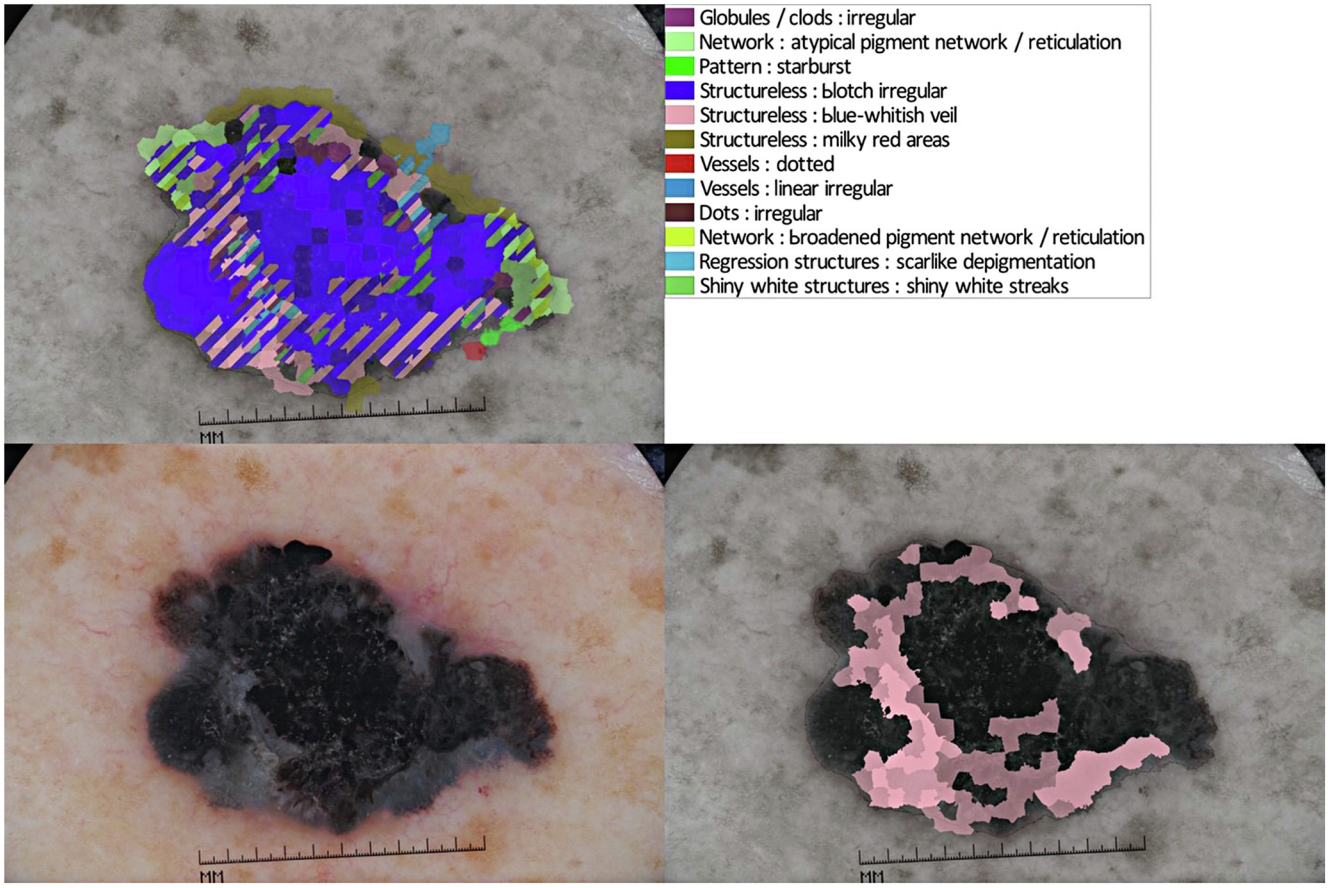
Yélamos O, Braun RP, Liopyris K, Wolner ZJ, Kerl K, Gerami P, et al. Dermoscopy and dermatopathology correlates of cutaneous neoplasms. J Am Acad Dermatol 2019;80:341–63.

Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans Med Imaging 1994;13:716–24.

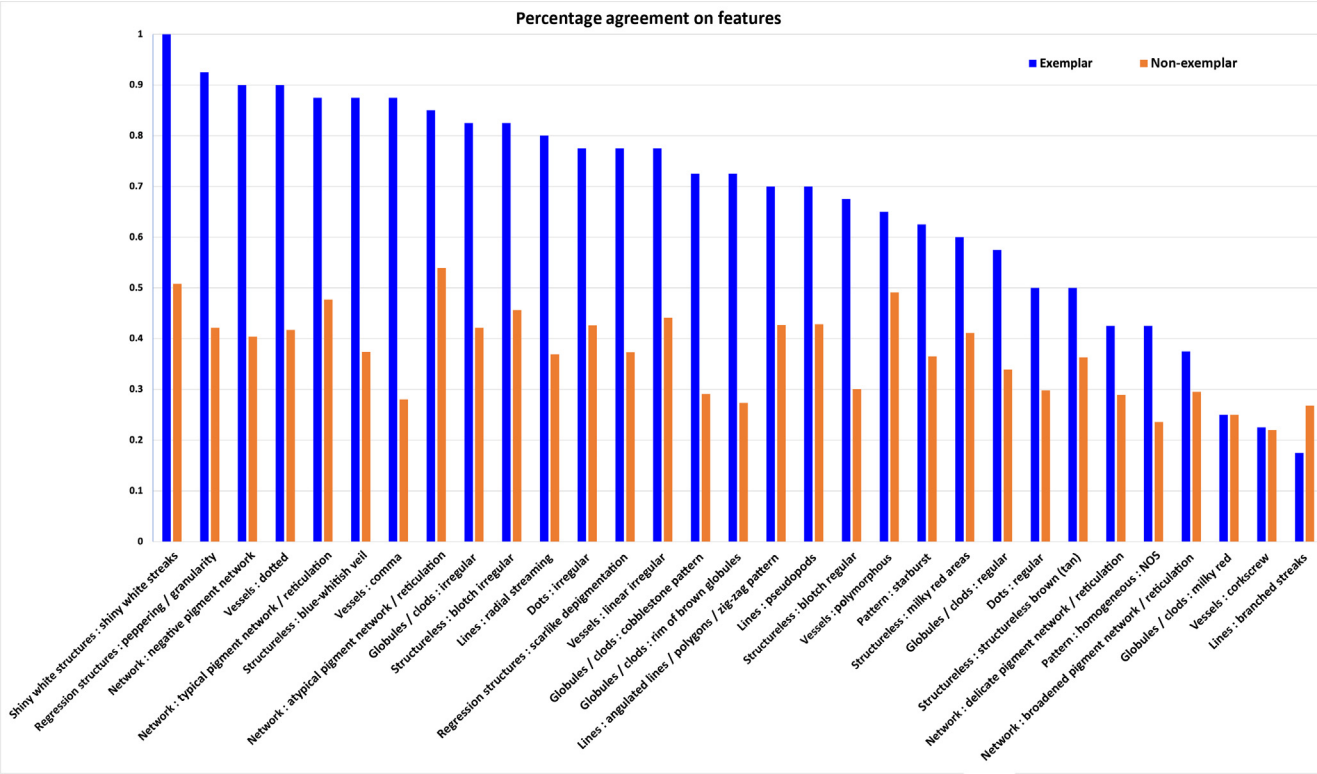
SUPPLEMENTARY RESULTS

Experts included in this study were five experts from the United States; three from Spain; two from Austria; two from Italy; and one expert from Australia, Chile, Colombia, France, Germany, Greece, Israel, and Switzerland. A total of 30% of experts were female (n = 6); all readers had >10 years of dermoscopy experience.

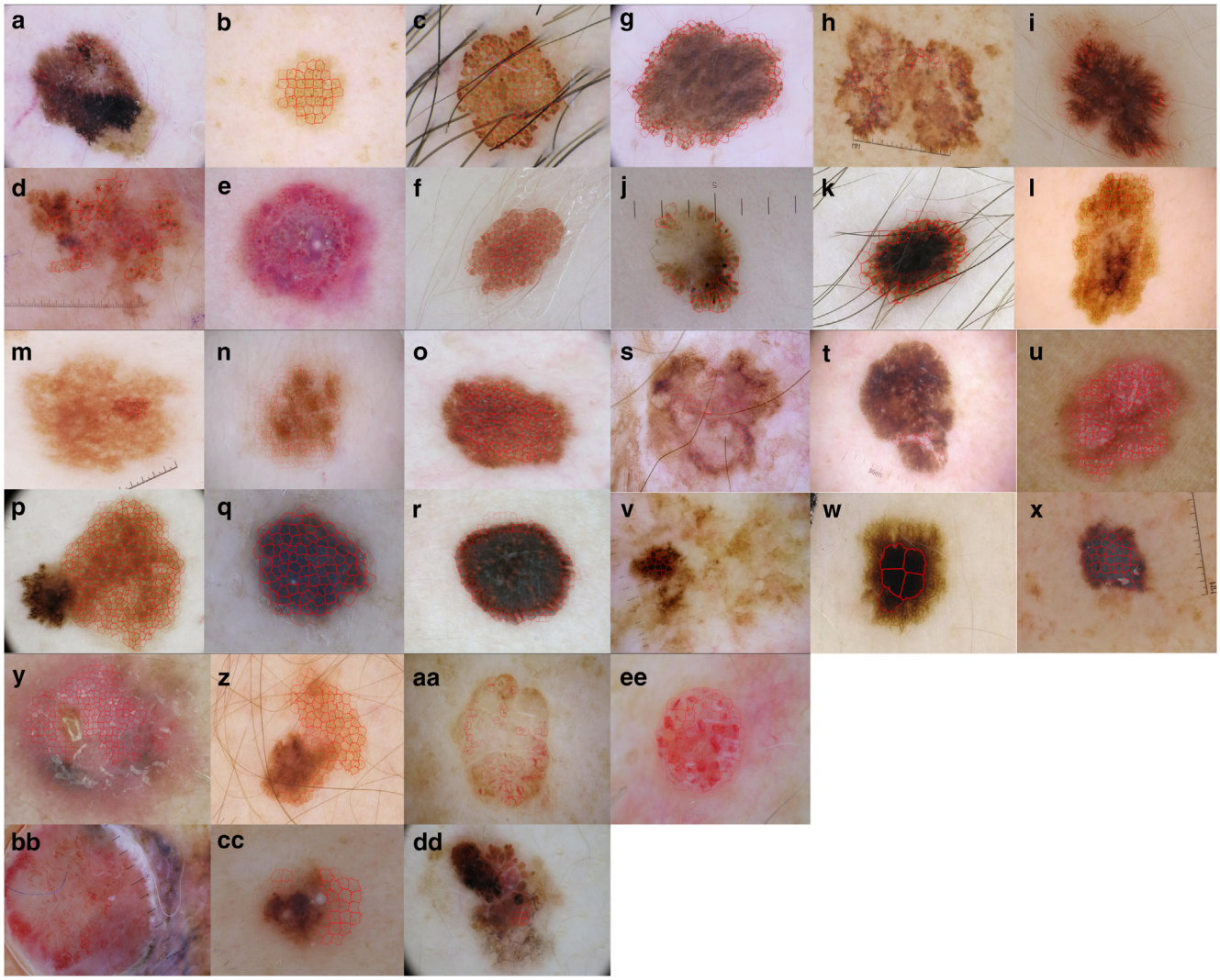
The median annotation time per image was 2:37 minutes (interquartile range = 1:25−5:11 minutes). Each reader annotated an average of 3.8 (SD = 2.4 per annotation, SD = 1.53 across images) features per lesion for a total of 4,507 feature markups. The average number of features annotated per image by the experts varied per diagnosis, from 2.86 (SD = 1.80) for nevi up to 4.63 (SD = 2.55) for melanomas ($P < 0.001$).



Supplementary Figure S1. Exemplar map of superpixel annotations. Example of an image included in the study (ISIC_0016128), displaying both examples of high-superpixel agreement among readers for the exemplar feature (structureless:blue–whitish veil, highlighted in pink in these annotations) and high Dice overlap among annotations between network:atypical pigment network/reticulation and network:broadened pigment network/reticulation.



Supplementary Figure S2. Percentage agreement on the presence of features across all 248 images on the lesion level. An agreement of 100% represents positive identification of the feature by all 20 experts for all times the feature was identified in an image; with blue, we see the agreement when the features occurred in exemplar images, and with orange, we see the agreement for the same features when they were identified in nonexemplar images.



Supplementary Figure S3. Exemplar images with a high level of agreement for each of the studied features. The areas of agreement are displayed with the outlines of superpixels, presented with different densities to highlight the areas of highest agreement. **(a)** ISIC_0046917, dots:irregular; (**b)** ISIC_0046364, dots:regular; (**c)** ISIC_0046462, globules/clods:cobblestone pattern; (**d)** ISIC_0016084, globules/clods:irregular; (**e)** ISIC_0023713, globules/clods:milky red globules; (**f)** ISIC_0046612, globules/clods:regular; **(g)** ISIC_0046642, globules/clods:rim of brown globule; (**h)** ISIC_0016123, lines:angulated lines/polygons/zig-zag; lines; (**i)** ISIC_0016094, branched streaks; (**j)** ISIC_0046615 lines:pseudopods; (**k)** ISIC_0047016, lines:radial streaming; (**l)** ISIC_0016114, network:atypical pigment network/reticulation; **(b)** Network:broadened pigment network/reticulation; (**n)** ISIC_0021052, network:delicate pigment network/reticulation; (**o)** ISIC_ 0046683, network:negative pigment network; (**p)** ISIC_0046622, network:typical pigment network/reticulation; (**q)** ISIC_0046606, pattern:homogeneous; (**r)** ISIC_0046591, pattern:starburst; (**s)** ISIC_0021271, regression structures:peppering/granularity; (**t**) ISIC_46879, regression structures:scarlike depigmentation; (**u)** ISIC_0046334, shiny white structures:shiny white streaks; (**v**) ISIC_0022661, structureless:blotch irregular; (**w)** ISIC_0016150, structureless:blotch regular; (**x)** ISIC_0016080, structureless:blue–whitish veil; (**y)** ISIC_004664, structureless:milky red areas; (**z)** ISIC_0046741, structureless:tan peripheral structureless areas; (**aa)** ISIC_0046789, vessels:comma; (**bb)** ISIC_0046682, vessels:corkscrew; (**cc)** ISIC_0016127, vessels:dotted; (**dd)** ISIC_0046361, vessels:linear irregular; and (**ee)** ISIC_0046890, vessels:polymorphous.

Supplementary Table S1. List of All Images Included in the Study with Their Respective Image IDs

ISIC Name	ISIC Image ID	Exemplar Feature for Which the Images Were Submitted
ISIC_0016081	589de95dd831136be37e0d4f	Network:broadened pigment network/reticulation
ISIC_0016092	58b0a360d831137d0a388356	Structureless:milky red areas
ISIC_0016127	58b0a372d831137d0a3884a1	Vessels:dotted
ISIC_0016131	58b0a374d831137d0a3884c5	Vessels:linear irregular
ISIC_0016137	58b0a377d831137d0a388507	Network:atypical pigment network/reticulation
ISIC_0016153	58b0a37dd831137d0a388597	Lines:angulated lines/polygons/zig-zag pattern
ISIC_0016159	58b0a380d831137d0a3885cd	Globules/clods:cobblestone pattern
ISIC_0021113	59e509ddd831136981ee66ae	Dots:regular
ISIC_0021271	59e50a0dd831136981ee6cb4	Regression structures:peppering/granularity
ISIC_0021713	59e50a98d831136981ee7d6a	Lines:angulated lines/polygons/zig-zag pattern
ISIC_0023234	59e50c82d831136981eeb733	Dots:regular
ISIC_0023252	59e50c8cd831136981eeb7ed	Network:atypical pigment network/reticulation
ISIC_0046222	5bf31fd31165972a676ac44a	Pattern:starburst
ISIC_0046224	5bf31fd31165972a676ac463	Lines:branched streaks
ISIC_0046240	5bf31fd71165972a676ac533	Vessels:linear irregular
ISIC_0046294	5bf31fe41165972a676ac80c	Network:delicate pigment network/reticulation
ISIC_0046316	5bf31fec1165972a676ac91b	Structureless:blotch regular
ISIC_0046338	5bf31ffa1165972a676aca41	Network:negative pigment network
ISIC_0046355	5bf320041165972a676acb78	Globules/clods:cobblestone pattern
ISIC_0046383	5bf320141165972a676acce5	Regression structures:peppering/granularity
ISIC_0046416	5bf320261165972a676ace90	Regression structures:scarlike depigmentation
ISIC_0046428	5bf3202d1165972a676acf23	Globules/clods:regular
ISIC_0046429	5bf3202d1165972a676acf2e	Pattern:homogeneous: NOS
ISIC_0046438	5bf320311165972a676acf97	Regression structures:scarlike depigmentation
ISIC_0046493	5bf3203d1165972a676ad22e	Lines:radial streaming
ISIC_0046529	5bf320441165972a676ad3d5	Globules/clods:regular
ISIC_0046546	5bf320491165972a676ad491	Pattern:homogeneous: NOS
ISIC_0046582	5bf320501165972a676ad645	Vessels:dotted
ISIC_0046597	5bf320551165972a676ad6ee	Structureless:structureless brown (tan)
ISIC_0046609	5bf320571165972a676ad77d	Shiny white structures:shiny white streaks
ISIC_0046613	5bf320571165972a676ad7ad	Network: typical pigment network/reticulation
ISIC_0046615	5bf320581165972a676ad7c3	Lines: pseudopods
ISIC_0046623	5bf320591165972a676ad81e	Structureless:blotch irregular
ISIC_0046631	5bf3205b1165972a676ad87a	Lines:radial streaming
ISIC_0046633	5bf3205b1165972a676ad891	Globules/clods:irregular
ISIC_0046635	5bf3205b1165972a676ad8aa	Pattern:starburst
ISIC_0046641	5bf3205d1165972a676ad8f6	Shiny white structures:shiny white streaks
ISIC_0046645	5bf3205d1165972a676ad928	Globules/clods:irregular
ISIC_0046659	5bf320601165972a676ad9d5	Dots:irregular
ISIC_0046661	5bf320621165972a676ad9f8	Vessels:corkscrew
ISIC_0046664	5bf320621165972a676ada19	Globules/clods:milky red
ISIC_0046671	5bf320641165972a676ada6a	Vessels:polymorphous
ISIC_0046674	5bf320651165972a676ada8b	Lines:branched streaks
ISIC_0046683	5bf320671165972a676adaee	Network:negative pigment network
ISIC_0046700	5bf3206d1165972a676adbb0	Vessels:comma
ISIC_0046745	5bf3207e1165972a676adda7	Network: typical pigment network/reticulation
ISIC_0046820	5bf320901165972a676ae135	Dots:irregular
ISIC_0046827	5bf320931165972a676ae199	Structureless:milky red areas
ISIC_0046834	5bf320961165972a676ae1f3	Vessels:corkscrew
ISIC_0046851	5bf3209c1165972a676ae2c2	Globules/clods: rim of brown globules
ISIC_0046852	5bf3209c1165972a676ae2ce	Globules/clods: rim of brown globules
ISIC_0046884	5bf320af1165972a676ae48d	Structureless:blue–whitish veil
ISIC_0046886	5bf320af1165972a676ae4a6	Globules/clods:milky red
ISIC_0046898	5bf320b81165972a676ae54f	Network:broadened pigment network/reticulation
ISIC_0046932	5bf320c61165972a676ae70d	Lines:pseudopods
ISIC_0046937	5bf320c91165972a676ae74d	Network:delicate pigment network/reticulation
ISIC_0046958	5bf320d51165972a676ae865	Structureless:structureless brown (tan)
ISIC_0046971	5bf320e11165972a676ae92d	Structureless:blotch regular

(*continued*)

Supplementary Table S1. Continued

ISIC Name	ISIC Image ID	Exemplar Feature for Which the Images Were Submitted
ISIC_0047002	5bf320ec1165972a676aeaa7	Vessels:polymorphous
ISIC_0047022	5bf320f41165972a676aeb96	Structureless:blue–whitish veil
ISIC_0047031	5bf320f61165972a676aec04	Structureless:blotch irregular
ISIC_0047032	5bf320f61165972a676aec13	Vessels:comma
ISIC_0016080	589de95dd831136be37e0d42	Structureless:blue–whitish veil
ISIC_0016082	589de95ed831136be37e0d58	Vessels:corkscrew
ISIC_0016084	589de95ed831136be37e0d6a	Globules/clods:irregular
ISIC_0016101	58b0a364d831137d0a3883b6	Globules/clods:milky red
ISIC_0016103	58b0a365d831137d0a3883c9	Network:atypical pigment network/reticulation
ISIC_0016105	58b0a366d831137d0a3883db	Pattern:homogeneous: NOS
ISIC_0016114	58b0a36bd831137d0a38842c	Dots:irregular
ISIC_0016125	58b0a371d831137d0a38848f	Dots:regular
ISIC_0016145	58b0a379d831137d0a38854f	Structureless:structureless brown (tan)
ISIC_0016165	5915aa64d83113352ace7a1d	Network:typical pigment network/reticulation
ISIC_0016175	5915aa76d83113352ace7a98	Regression structures:scarlike depigmentation
ISIC_0021901	59e50ad4d831136981ee8496	Structureless:milky red areas
ISIC_0022661	59e50bcbd831136981eea176	Structureless:blotch irregular
ISIC_0046302	5bf31fe61165972a676ac871	Structureless:milky red areas
ISIC_0046322	5bf31fef1165972a676ac969	Network:negative pigment network
ISIC_0046336	5bf31ff91165972a676aca24	Vessels:linear irregular
ISIC_0046343	5bf31ffd1165972a676aca96	Vessels:polymorphous
ISIC_0046361	5bf320081165972a676acbd4	Regression structures:peppering/granularity
ISIC_0046372	5bf3200f1165972a676acc55	Dots:irregular
ISIC_0046385	5bf320151165972a676accfd	Network:delicate pigment network/reticulation
ISIC_0046397	5bf3201c1165972a676acd92	Structureless:blotch regular
ISIC_0046422	5bf3202a1165972a676acee1	Lines:angulated lines/polygons/zig-zag pattern
ISIC_0046441	5bf320321165972a676acfc7	Vessels:dotted
ISIC_0046449	5bf320341165972a676ad02b	Structureless:blue–whitish veil
ISIC_0046504	5bf3203e1165972a676ad2af	Lines:angulated lines/polygons/zig-zag pattern
ISIC_0046557	5bf3204b1165972a676ad514	Regression structures:peppering/granularity
ISIC_0046560	5bf3204c1165972a676ad539	Network:atypical pigment network/reticulation
ISIC_0046596	5bf320541165972a676ad6e3	Network:negative pigment network
ISIC_0046606	5bf320571165972a676ad75c	Pattern:hHomogeneous:NOS
ISIC_0046617	5bf320581165972a676ad7d9	Globules/clods:rim of brown globules
ISIC_0046618	5bf320581165972a676ad7e4	Network:broadened pigment network/reticulation
ISIC_0046626	5bf3205a1165972a676ad83f	Vessels:corkscrew
ISIC_0046629	5bf3205a1165972a676ad860	Globules/clods:cobblestone pattern
ISIC_0046640	5bf3205c1165972a676ad8e8	Regression structures:scarlike depigmentation
ISIC_0046665	5bf320621165972a676ada24	Vessels:dotted
ISIC_0046666	5bf320621165972a676ada2f	Vessels:comma
ISIC_0046682	5bf320671165972a676adae3	Vessels:polymorphous
ISIC_0046684	5bf320681165972a676adaf9	Structureless:structureless brown (tan)
ISIC_0046693	5bf3206c1165972a676adb5f	Lines:branched streaks
ISIC_0046724	5bf320761165972a676adcc0	Lines:pseudopods
ISIC_0046725	5bf320761165972a676adccb	Shiny white structures:shiny white streaks
ISIC_0046732	5bf320791165972a676add18	Globules/clods:rim of brown globules
ISIC_0046740	5bf3207d1165972a676add70	Pattern:starburst
ISIC_0046786	5bf320871165972a676adfa1	Globules/clods:regular
ISIC_0046798	5bf3208b1165972a676ae02b	Vessels:comma
ISIC_0046862	5bf320a01165972a676ae35e	Lines:branched streaks
ISIC_0046864	5bf320a11165972a676ae37c	Pattern:starburst
ISIC_0046870	5bf320a41165972a676ae3ca	Structureless:blotch irregular
ISIC_0046880	5bf320ab1165972a676ae446	Network:typical pigment network/reticulation
ISIC_0046903	5bf320ba1165972a676ae59d	Globules/clods:regular
ISIC_0046908	5bf320bd1165972a676ae5e7	Shiny white structures:shiny white streaks
ISIC_0046918	5bf320c11165972a676ae668	Globules/clods:irregular
ISIC_0046949	5bf320cf1165972a676ae7ee	Globules/clods:cobblestone pattern
ISIC_0046951	5bf320d01165972a676ae80c	Dots:regular

(*continued*)

Supplementary Table S1. Continued

ISIC Name	ISIC Image ID	Exemplar Feature for Which the Images Were Submitted
ISIC_0046965	5bf320dc1165972a676ae8d0	Network:broadened pigment network/reticulation
ISIC_0046967	5bf320de1165972a676ae8f5	Lines:pseudopods
ISIC_0046969	5bf320df1165972a676ae913	Structureless:blotch regular
ISIC_0046980	5bf320e61165972a676ae997	Network:delicate pigment network/reticulation
ISIC_0047004	5bf320ed1165972a676aeac5	Lines:radial streaming
ISIC_0047008	5bf320ee1165972a676aeaf1	Vessels:linear irregular
ISIC_0047016	5bf320f21165972a676aeb4d	Lines:radial streaming
ISIC_0047033	5bf320f71165972a676aec22	Globules/clods:milky red
ISIC_0016085	589de95fd831136be37e0d73	Structureless:blotch irregular
ISIC_0016162	58b0a382d831137d0a3885e8	Globules/clods:milky red
ISIC_0016168	5915aa6dd83113352ace7a4b	Shiny white structures:shiny white streaks
ISIC_0021092	59e509d7d831136981ee65e5	Network:broadened pigment network/reticulation
ISIC_0021221	59e509fdd831136981ee6ad3	Globules/clods:cobblestone pattern
ISIC_0021442	59e50a44d831136981ee732f	Lines:angulated lines/polygons/zig-zag pattern
ISIC_0021616	59e50a7cd831136981ee79c5	Network:atypical pigment network/reticulation
ISIC_0022328	59e50b5fd831136981ee94cb	Structureless:blue–whitish veil
ISIC_0023162	59e50c6cd831136981eeb480	Vessels:corkscrew
ISIC_0023713	59e50d52d831136981eec9d9	Globules/clods:milky red
ISIC_0046223	5bf31fd31165972a676ac455	Pattern:starburst
ISIC_0046225	5bf31fd31165972a676ac46f	Lines:branched streaks
ISIC_0046242	5bf31fd71165972a676ac54d	Regression structures:scarlike depigmentation
ISIC_0046305	5bf31fe71165972a676ac896	Pattern:homogeneous: NOS
ISIC_0046324	5bf31ff01165972a676ac982	Network:typical pigment network/reticulation
ISIC_0046331	5bf31ff51165972a676ac9de	Shiny white structures:shiny white streaks
ISIC_0046332	5bf31ff51165972a676ac9e9	Network:broadened pigment network/reticulation
ISIC_0046335	5bf31ff71165972a676aca16	Dots:regular
ISIC_0046340	5bf31ffb1165972a676aca61	Vessels:corkscrew
ISIC_0046344	5bf31ffe1165972a676acaab	Globules/clods:irregular
ISIC_0046345	5bf31ffe1165972a676acabe	Vessels:linear irregular
ISIC_0046358	5bf320061165972a676acbac	Dots:regular
ISIC_0046409	5bf320231165972a676ace30	Lines:branched streaks
ISIC_0046427	5bf3202c1165972a676acf18	Network:atypical pigment network/reticulation
ISIC_0046443	5bf320321165972a676acfdd	Network:negative pigment network
ISIC_0046450	5bf320341165972a676ad036	Structureless:blue–whitish veil
ISIC_0046454	5bf320351165972a676ad067	Globules/clods:rim of brown globules
ISIC_0046471	5bf320391165972a676ad12d	Structureless:structureless brown (tan)
ISIC_0046495	5bf3203d1165972a676ad244	Structureless:blotch irregular
ISIC_0046548	5bf320491165972a676ad4a9	Dots:irregular
ISIC_0046608	5bf320571165972a676ad772	Structureless:blotch regular
ISIC_0046612	5bf320571165972a676ad79e	Globules/clods:regular
ISIC_0046628	5bf3205a1165972a676ad855	Globules/clods:cobblestone pattern
ISIC_0046639	5bf3205c1165972a676ad8dd	Regression structures:peppering/granularity
ISIC_0046642	5bf3205d1165972a676ad902	Globules/clods:rim of brown globules
ISIC_0046647	5bf3205e1165972a676ad941	Dots:irregular
ISIC_0046667	5bf320631165972a676ada3c	Structureless:milky red areas
ISIC_0046669	5bf320631165972a676ada54	Regression structures:peppering/granularity
ISIC_0046670	5bf320631165972a676ada5f	Vessels:dotted
ISIC_0046706	5bf3206e1165972a676adbf6	Structureless:milky red areas
ISIC_0046741	5bf3207d1165972a676add7b	Structureless:structureless brown (tan)
ISIC_0046784	5bf320871165972a676adf85	Structureless:blotch regular
ISIC_0046789	5bf320881165972a676adfc8	Vessels:polymorphous
ISIC_0046811	5bf3208e1165972a676ae0bd	Network:delicate pigment network/reticulation
ISIC_0046812	5bf3208e1165972a676ae0cb	Vessels:comma
ISIC_0046823	5bf320921165972a676ae162	Vessels:comma
ISIC_0046828	5bf320931165972a676ae1a5	Lines:angulated lines/polygons/zig-zag pattern
ISIC_0046848	5bf3209b1165972a676ae29c	Network:typical pigment network/reticulation
ISIC_0046861	5bf320a01165972a676ae351	Lines:pseudopods
ISIC_0046865	5bf320a21165972a676ae38e	Pattern:starburst

(*continued*)

Supplementary Table S1. Continued

ISIC Name	ISIC Image ID	Exemplar Feature for Which the Images Were Submitted
ISIC_0046879	5bf320aa1165972a676ae433	Regression structures:scarlike depigmentation
ISIC_0046891	5bf320b11165972a676ae4e9	Vessels:dotted
ISIC_0046900	5bf320b81165972a676ae56d	Network:negative pigment network
ISIC_0046917	5bf320c11165972a676ae659	Globules/clods:irregular
ISIC_0046921	5bf320c31165972a676ae691	Lines:radial streaming
ISIC_0046966	5bf320dd1165972a676ae8e3	Pattern:homogeneous:NOS
ISIC_0046977	5bf320e41165972a676ae973	Network:delicate pigment network/reticulation
ISIC_0046983	5bf320e71165972a676ae9c0	Vessels:linear irregular
ISIC_0046985	5bf320e81165972a676ae9d6	Lines:radial streaming
ISIC_0047018	5bf320f21165972a676aeb66	Vessels:polymorphous
ISIC_0047019	5bf320f31165972a676aeb71	Globules/clods:regular
ISIC_0047023	5bf320f41165972a676aeba4	Lines:pseudopods
ISIC_0016094	58b0a361d831137d0a38836c	Network:atypical pigment network/reticulation
ISIC_0016112	58b0a36ad831137d0a38841a	Lines:angulated lines/polygons/zig-zag pattern
ISIC_0016123	58b0a370d831137d0a38847d	Regression structures:peppering/granularity
ISIC_0016128	58b0a372d831137d0a3884aa	Structureless: Blue-whitish veil
ISIC_0016133	58b0a375d831137d0a3884d7	Structureless:blotch irregular
ISIC_0016139	58b0a377d831137d0a388519	Vessels:polymorphous
ISIC_0016143	58b0a379d831137d0a38853d	Lines:pseudopods
ISIC_0016150	58b0a37cd831137d0a38857c	Structureless:blotch regular
ISIC_0016161	58b0a382d831137d0a3885df	Globules/clods:milky red
ISIC_0021052	59e509ccd831136981ee6459	Network:delicate pigment network/reticulation
ISIC_0021894	59e50ad1d831136981ee844f	Pattern:homogeneous: NOS
ISIC_0023588	59e50d2dd831136981eec519	Vessels:corkscrew
ISIC_0023593	59e50d31d831136981eec54d	Dots:irregular
ISIC_0023811	59e50d71d831136981eecd89	Lines:radial streaming
ISIC_0046261	5bf31fdc1165972a676ac64f	Pattern:starburst
ISIC_0046315	5bf31fec1165972a676ac910	Network:atypical pigment network/reticulation
ISIC_0046317	5bf31fed1165972a676ac92e	Vessels:dotted
ISIC_0046323	5bf31ff01165972a676ac977	Network:negative pigment network
ISIC_0046334	5bf31ff61165972a676aca07	Shiny white structures:shiny white streaks
ISIC_0046351	5bf320031165972a676acb40	Globules/clods:cobblestone pattern
ISIC_0046353	5bf320031165972a676acb59	Vessels:comma
ISIC_0046364	5bf320091165972a676acbfa	Globules/clods:regular
ISIC_0046394	5bf3201b1165972a676acd6d	Regression structures:peppering/granularity
ISIC_0046405	5bf320201165972a676acdf7	Globules/clods:rim of brown globules
ISIC_0046410	5bf320231165972a676ace3b	Lines:branched streaks
ISIC_0046414	5bf320261165972a676ace71	Shiny white structures:shiny white streaks
ISIC_0046458	5bf320361165972a676ad097	Structureless:blue–whitish veil
ISIC_0046462	5bf320371165972a676ad0c7	Globules/clods:cobblestone pattern
ISIC_0046499	5bf3203e1165972a676ad277	Network:delicate pigment network/reticulation
ISIC_0046526	5bf320431165972a676ad3b0	Structureless:milky red areas
ISIC_0046537	5bf320461165972a676ad42d	Globules/clods:milky red
ISIC_0046571	5bf3204e1165972a676ad5c4	Vessels:dotted
ISIC_0046572	5bf3204f1165972a676ad5cf	Pattern:homogeneous:NOS
ISIC_0046591	5bf320521165972a676ad6ac	Pattern:starburst
ISIC_0046593	5bf320531165972a676ad6c2	Vessels:linear irregular
ISIC_0046598	5bf320551165972a676ad6fc	Structureless:structureless brown (tan)
ISIC_0046607	5bf320571165972a676ad767	Structureless:blotch regular
ISIC_0046622	5bf320591165972a676ad812	Network:broadened pigment network/reticulation
ISIC_0046643	5bf3205d1165972a676ad90d	Vessels:linear irregular
ISIC_0046644	5bf3205d1165972a676ad918	Network:negative pigment network
ISIC_0046662	5bf320621165972a676ada03	Vessels:corkscrew
ISIC_0046672	5bf320651165972a676ada75	Regression structures:scarlike depigmentation
ISIC_0046673	5bf320651165972a676ada80	Lines:branched streaks
ISIC_0046678	5bf320661165972a676adab7	Globules/clods:irregular
ISIC_0046681	5bf320671165972a676adad8	Structureless:structureless brown (tan)
ISIC_0046712	5bf320701165972a676adc3c	Dots:irregular

(*continued*)

Supplementary Table S1. Continued

ISIC Name	ISIC Image ID	Exemplar Feature for Which the Images Were Submitted
ISIC_0046729	5bf320771165972a676adcf7	Dots:regular
ISIC_0046744	5bf3207e1165972a676add9c	Network:typical pigment network/reticulation
ISIC_0046756	5bf320801165972a676ade23	Lines:radial streaming
ISIC_0046773	5bf320841165972a676adef8	Structureless:blotch irregular
ISIC_0046797	5bf3208a1165972a676ae020	Vessels:comma
ISIC_0046866	5bf320a21165972a676ae39a	Dots:regular
ISIC_0046871	5bf320a41165972a676ae3d5	Network:typical pigment network/reticulation
ISIC_0046882	5bf320ad1165972a676ae46b	Regression structures:scarlike depigmentation
ISIC_0046890	5bf320b11165972a676ae4da	Vessels:polymorphous
ISIC_0046922	5bf320c41165972a676ae69c	Globules/clods:rim of brown globules
ISIC_0046943	5bf320ca1165972a676ae790	Network:broadened pigment network/reticulation
ISIC_0046955	5bf320d31165972a676ae844	Globules/clods:irregular
ISIC_0046956	5bf320d41165972a676ae84f	Structureless:milky red areas
ISIC_0046993	5bf320ea1165972a676aea3c	Globules/clods:regular
ISIC_0047003	5bf320ed1165972a676aeab6	Lines:pseudopods
ISIC_0047024	5bf320f41165972a676aebb0	Lines:angulated lines/polygons/zig-zag pattern

Abbreviations: ID, identification; ISIC, International Skin Imaging Collaboration; NOS, not otherwise specified.

The exemplar dermoscopic features of the images for which they were submitted by the experts can be found in the table. All images are available at www.isic-archive.com.

Supplementary Table S2. Dermoscopic Features Specific for Melanocytic Lesions (Nevi and Melanomas) Included in Our Study, the Total Number of Observations (on the Lesion Level) of These Features (Total Number of Observations), the Number of Images in Which They Were Observed (In Images), and the Respected Agreement They Yielded

Dermoscopic Feature	Total Number of Observations	In Images	Orphan Observations	Lesions ≥ 40% Agreement	Lesions ≥ 60% Agreement	Lesions ≥ 80% Agreement	Lesions with 100% Agreement
Dots:irregular	268	124	48	76	44	18	6
Dots:regular	102	65	37	28	7	2	0
Globules/clods:cobblestone pattern	60	29	14	15	9	5	2
Globules/clods:irregular	319	150	67	83	49	27	10
Globules/clods:milky red	40	31	24	7	2	0	0
Globules/clods:regular	103	57	28	29	14	3	0
Globules/clods:rim of brown globules	58	29	18	11	8	6	4
Lines:angulated lines/polygons/ zig-zag pattern	72	30	9	21	16	5	0
Lines:branched streaks	82	61	43	18	3	0	0
Lines:pseudopods	116	50	19	31	23	11	1
Lines:radial streaming	135	66	33	33	21	12	3
Network:atypical pigment network/reticulation	344	128	32	96	67	39	14
Network:broadened pigment network/reticulation	130	86	51	35	9	0	0
Network:delicate pigment Network/reticulation	124	81	55	26	12	5	0
Network:negative pigment network	124	54	22	32	19	13	6
Network:typical pigment network/ reticulation	245	101	34	67	39	26	12
Pattern:homogeneous:NOS	88	71	57	14	2	1	0
Pattern:starburst	53	24	10	14	8	5	2
Regression structures:peppering/ granularity	249	114	48	66	35	23	11
Regression structures:scarlike depigmentation	161	81	42	39	24	11	6
Shiny white structures:shiny white streaks	218	82	24	58	40	25	13
Structureless:blotch irregular	227	98	44	54	37	27	11
Structureless:blotch regular	71	39	24	15	9	6	2
Structureless:blue–whitish veil	183	93	39	54	27	6	3
Structureless:milky red areas	169	80	32	48	25	15	1
Structureless:structureless brown (tan)	250	138	62	76	28	7	1
Vessels:comma	63	28	16	12	10	8	5
Vessels:corkscrew	20	17	14	3	0	0	0
Vessels:dotted	147	63	26	37	25	15	7
Vessels:linear irregular	143	61	22	39	25	16	2
Vessels:polymorphous	143	54	14	40	28	16	5

Orphan observations indicate observations by single readers.

Supplementary Table S3. Terminology Glossary

Metaphoric Terminology	Definition	Descriptive Terminology	Abbreviation
Dots:regular	Dots clustered at the center of the lesion or located on the network lines or in the hole of the network (also called target network)	Target: dots, brown, central (in the center of hypopigmented spaces between reticular lines)	Regular dots
Dots:irregular	Any distribution of dots other than dots as described for regular dots		Irregular dots
Globules/clods:cobblestone pattern	Polygonal globules symmetrically distributed throughout the lesion	Clods, brown or skin colored, large, and polygonal	Cobblestone pattern
Globules/clods:irregular	Globules with variability in color, size, shape, or spacing and distributed in an asymmetric/ disorganized fashion		Irregular globules
Globules/clods:milky red		Clods, pink, and small	Milky-red globules
Globules/clods:regular	Globules with minimal variability in their color, size, and shape located in the center of a lesion with the surrounding network or along the perimeter or throughout the entire lesion	Clods, small, round, or oval	Regular globules
Globules/clods:rim of brown globules	Globules distributed at the periphery of the lesion	Clods, brown, circumferential	Rim of brown globules
Lines:angulated lines/ polygons/zig-zag	Gray–brown lines that are connected at an angle or coalescing to form polygons	Lines, angulated, or polygonal (nonfacial skin)	Angulated lines
Lines:branched streaks	Atypical network with broken/interrupted lines and incomplete connections		Branched streaks
Lines:pseudopods	Bulbous and often kinked projections seen at the lesion edge, either directly associated with a network or solid tumor border		Pseudopods
Lines:radial streaming	Radial linear extensions at the lesion edge	Lines, radial and segmental	Radial streaming
Network:atypical pigment network/reticulation	Network with increased variability in line color, thickness, and spacing. Gray color to lines or disorganized distribution	Lines, reticular and thick, or reticular lines that vary in color	Atypical network
Network:broadened pigment network/reticulation	Widening of the network lines	Lines, reticular, and thick	Broadened network
Network: delicate pigment network/reticulation	Fine thin network	Lines, reticular, and thin	Delicate network
Network: typical pigment network/reticulation	Network with minimal variability in the color, thickness, and spacing of the lines; symmetrically distributed	Lines, reticular	Typical network
Network:negative pigment network/reticulation	Serpiginous interconnecting broadened hypopigmented lines that surround elongated and curvilinear brown structures	Lines, reticular, and hypopigmented around brown clods	Negative network
Patterns:starburst pattern	This pattern consists of tiered peripheral globules, pseudopods, or streaks (or a combination of them), located around the entire perimeter of the lesion	Pseudopods, circumferential or lines, radial, circumferential	Starburst pattern
Patterns:homogeneous pattern	A pattern lacking any definable pigment structures, also known as structureless pattern	Structureless, any color	Homogeneous pattern
Regression structures:peppering/ granularity	Consists of fine dots with a blue–gray color	Dots, gray	Peppering/granularity
Regression structures:scar-like depigmentation	Area of white that is whiter than surrounding normal-appearing skin (true scarring); it should not be confused with hypopigmentation or depigmentation caused by simple loss of melanin; shiny white structures and blood vessels are not seen in areas of regression	Structureless zone, white	Scar-like depigmentation
Shiny white structures:shiny white streaks	Short discrete white lines oriented parallel and orthogonal (perpendicular) to each other seen only under polarized dermoscopy	Lines, white, perpendicular	Shiny white streaks
Structureless:blue–whitish veil	A raised/palpable blotch of blue hue with an overlying whitish ground-glass haze	Structureless zone, blue	Blue–whitish veil
Structureless:blotch Regular	One blotch within the center of the lesion and surrounded by network		Regular blotch
Structureless:blotch irregular	More than one blotch or a blotch that is located off center		Irregular blotch

(*continued*)

Supplementary Table S3. Continued

Metaphoric Terminology	Definition	Descriptive Terminology	Abbreviation
Structureless:tan (brown) peripheral structureless areas		Structureless, brown (tan), eccentric	Tan peripheral structureless areas
Structureless: milky-red areas	Milky-white appearance or pinkish structureless areas (strawberry and ice cream like), consisting of a red vascular blush with no specific distinguishable vessels		Milky-red areas
Vessel morphology, monomorphous:dots	Tiny pinpoint vessels		Dotted vessels
Vessel morphology, monomorphous:comma	Linear, curved, and short vessels	Curved	Comma vessels
Vessel morphology, monomorphous:corkscrew	Twisted looped vessels with bends twisted along a central axis	Helical	Corkscrew vessels
Vessel morphology, monomorphous:linear irregular/serpentine	Linear, curved, or serpentine vessels,	Serpentine	Linear irregular vessels
Vessel morphology:polymorphous			Polymorphous vessels

Dermoscopy has an extended terminology and two different approaches to the names of the features: metaphoric and descriptive. In the table, we provide the feature names and their correspondence between metaphoric and descriptive terminology, their definitions, and the abbreviations that are used in the manuscript. The table was adapted and modified with permission from Kittler et al. (2016).

Supplementary Table S4. Total Number of Superpixels Annotated for Each Feature by 20 Readers (Including Overlapping Superpixels) across 248 Images

Dermoscopic Feature	Total Number of Superpixels Annotated	3RA %	Full − 5RA %
Dots:irregular	4,034	9.20	0.20
Dots:regular	2,011	8.65	0.00
Globules/clods:cobblestone pattern	2,933	40.03	14.63
Globules/clods:irregular	6,332	18.62	1.14
Globules/clods:milky red	1,246	2.65	0.00
Globules/clods:regular	4,280	22.87	0.00
Globules/clods:rim of brown globules	1,133	31.24	10.86
Lines:angulated lines/polygons/zig-zag pattern	1,688	12.14	0.00
Lines:branched streaks	1,331	0.90	0.00
Lines:pseudopods	884	18.33	0.11
Lines:radial streaming	2,395	19.54	1.63
Network:atypical pigment network/reticulation	12,629	15.54	1.61
Network:broadened pigment network/reticulation	3,503	1.77	0.00
Network:delicate pigment network/reticulation	4,045	1.43	0.00
Network:negative pigment network	3,095	24.07	5.27
Network:typical pigment network/reticulation	10,551	39.42	11.88
Pattern:homogeneous: NOS	8,615	2.44	0.00
Pattern:starburst	2,503	32.68	3.36
Regression structures:peppering/granularity	6,697	28.15	3.82
Regression structures:scarlike depigmentation	4,442	14.77	1.06
Shiny white structures:shiny white streaks	4,228	27.08	4.99
Structureless:blotch irregular	6,613	17.51	2.37
Structureless:blotch regular	1,432	19.27	2.86
Structureless:blue–whitish veil	5,259	18.05	1.54
Structureless:milky red areas	5,980	12.61	0.00
Structureless:structureless brown (tan)	7,612	7.45	0.37
Vessels:comma	1,303	12.97	1.77
Vessels:corkscrew	161	0.00	0.00
Vessels:dotted	3,345	37.73	5.83
Vessels:linear irregular	3,107	14.00	0.68
Vessels:polymorphous	6,636	26.19	2.09

Abbreviations: NOS, Not otherwise specified; 3RA, 3-Reader agreement; 5RA, 5-Reader agreement.

The percentage of agreement ≥60% (at least three of five readers annotating the same region/superpixels for the same feature) can be seen under 3RA %, whereas under Full − 5RA%, we can see the percentage of absolute agreement for each feature (five of five readers annotating the same region/superpixels for the same feature).