RESEARCH ON MELANOCITIC SKIN LESION INFOBASE ENLARGEMENT – NEW FACTS AND CONCEPTS

DALIA CHRZANOWSKA[1], RAFAŁ NIEMIEC[2] AND ZDZISŁAW HIPPE[2]

[1] *Faculty of Medicine, UITM*
mjr. H. Sucharskiego 2, 25-225 Rzeszów, Poland

[2] *Faculty of Applied Computer Science, UITM*
mjr. H. Sucharskiego 2, 25-225 Rzeszów, Poland

(received: 23 August 2017; revised: 22 September 2017;
accepted: 29 September 2017; published online: 12 October 2017)

Abstract: The melanocytic skin lesion infobase, available at http://synthesis.melanoma.pl (also http://synteza.melanoma.pl, in Polish; referred to as **INP**) is currently undergoing a complete modification of the way in which (i) the internal synthesis algorithms and (ii) the classification of lesions are performed. We investigated 29 new real images of melanocytic skin lesions, focusing on how humans perform classification based on experience. In conclusion we suggest to add a new color – connected with the depth of a lesion – to the **K term** of Asymmetry (**A**), Border (**B**) and linear Combination of colors and structures (**K**) method (referred to as **ABK**).

Keywords: melanoma, ABK, attributes importance

DOI: https://doi.org/10.17466/tq2017/21.4/z

1. Infobase enlargement

We added 29 new **cases** (Figure 1) to the **INP** database, described by **A**symmetry, **B**order and *modified* **K term** (**ABK** method).

The cases were based on the visual attribute-value recognition of real lesion images[1] from [1, 2] The modification of the **K term** consisted in addition of a new attribute – the gray color. Gray was diagnosed in 14 cases. New cases were classified into three classes: (i) *melanoma malignum*, (ii) *lentigo*, (iii) *neavus*. 22 cases were classified to the *melanoma malignum* class, one case to the *lentigo* class and 6 cases to the *neavus* class. The cases formed an unbalanced set. A new color was identified – gray. While it was not blue-gray, but the gray color connected

1. With thanks to Sunil Kumar Iha for the help to get the images.

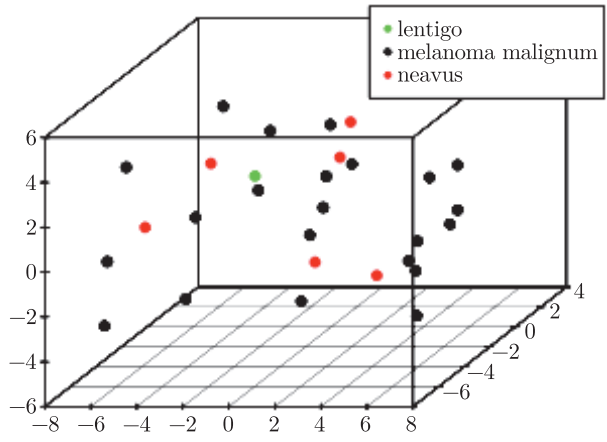

Figure 1. New cases (29) photographed using non-linear multidimensional scaling ($14 \to 3$, while maintaining relative city-block distances between cases) using the Kruskal method; recognized and evaluated by a dermatologist-venereologist

with the *depth* of a lesion, we checked experimentally the validity of its inclusion in the **K term**.

2. Importance of gray color

The set of cases was equalized to the majority class by sampling with replacement minor subset cases. The singleton *lentigo* was removed. All the important attributes [3] were selected from: (i) **K term** attributes and (ii) **ABK** attributes. To this end, the Boruta method [4] was used, with p-value $= 0.001$, without a Bonferroni correction and where *tentative* attributes were treated as *irrelevant* after 1000 of iterations. Boruta repeatedly does the two-sided equality test for every attribute rank with a highest ranked shadow (a random value-permutation attribute from predictors) and removes unimportant while marking important ones in every iteration. The Z-score of an average reduction of the 500 tree forest [5] accuracy was used as the source of ranks. Therefore, it was crucial to balance the dataset with reasonable cases. The procedure (balance and selection) was repeated 30 times. The quotients of *times* selected as *important* for the number of trials (referred to as *importance*) for: (i) **K term** and (ii) **ABK** attributes, including a standard error (in confidence intervals 0.95; 0.999) are shown in Table 1.

Gray was important in $15/30_{\pm 0.19^{95\%};\, 0.33^{99.9\%}}$ trials, while taking into consideration the **K term** alone and in $26/30_{\pm 0.13^{95\%};\, 0.23^{99.9\%}}$ trials when all **ABK** attributes were present. Moreover, the gray importance – as the only one – differs when using the Wilcoxon matched-pairs signed rank test. The attribute was added to the **K term**, marked as *tentative* for further experiments (with more data). This experiment showed which attributes were important – with an ad-hoc importance measure – in medical discrimination of melanoma malignum and neavus, when based on the doctor's experience.

Table 1. Attribute importance (relative to medical classification) of (i) **K term** and (ii) **ABK** attributes

K term		Wilcoxon pair test (p-value)	**ABK**	
ATTRIBUTE	IMPORTANCE		ATTRIBUTE	IMPORTANCE
			Asymmetry	$18/30_{\pm 0.18}$
			Border	$1_{\pm 0}$
White	$12/30_{\pm 0.19;\,0.33}$	0.3014	White	$8/30_{\pm 0.17;\,0.30}$
Gray	$15/30_{\pm 0.19;\,0.33}$	0.0050	Gray	$26/30_{\pm 0.13;\,0.23}$
Blue	$0_{\pm 0}$	–	Blue	$0_{\pm 0}$
Dark brown	$0_{\pm 0}$	–	Dark brown	$0_{\pm 0}$
Light brown	$0_{\pm 0}$	–	Light brown	$0_{\pm 0}$
Black	$21/30_{\pm 0.17;\,0.30}$	0.5941	Black	$23/30_{\pm 0.16;\,0.28}$
Red	$0_{\pm 0}$	–	Red	$0_{\pm 0}$
Pigment globules	$1_{\pm 0}$	1	Pigment globules	$29/30_{\pm 0.07;\,0.12}$
Pigment dots	$1_{\pm 0}$	–	Pigment dots	$1_{\pm 0}$
Structureless area	$25/30_{\pm 0.14;\,0.25}$	0.5297	Structureless area	$27/30_{\pm 0.11;\,0.20}$
Branched streaks	$1_{\pm 0}$	–	Branched streaks	$1_{\pm 0}$
Pigment network	$15/30_{\pm 0.19;\,0.33}$	0.6834	Pigment network	$17/30_{\pm 0.19;\,0.33}$

3. Conclusions

We added a new important attribute to the **K term** of the **ABK** method and 29 new cases to the melanocitic skin lesion **INP** infobase. The next step will be to develop methods to transfer the human experience to the expert system, *i.e.* to isolate the experience from the acquired knowledge and then transfer it – separately – to the expert system. To this end, we plan to work with at least three different dermatologists-venereologists. We will analyze *their* focus on a specific lesion phenomena in a process of discrimination and compare their results with the results of: (i) the protocol that was the base of their acquired knowledge (medical training) and (ii) histopathological lesion analysis (*a.k.a.* a truth).

References

[1] Celebi M E, Wan Q, Hwang S, Iyatomi H and Schaefer G 2013 *Skin Research and Technology* **19** (1) 252

[2] Celebi M E, Iyatomi H, Schaefer G and Stoecher W V 2009 *Computerized Medical Imaging and Graphics* **33** (2) 148

[3] Kohavi R and John G 1997 *Artificial Intelligence Journal, special issue on relevance* **97** (1-2) 273

[4] Rudnicki W R, Kierczak M, Koronacki J and Komorowski J 2006 *International Conference on Rough Sets and Current Trends in Computing*, Springer 557

[5] Breiman L 2001 *Machine Learning* **45.1** 5