

10-2020

Why Number of Color Difference Works Better In Detecting Melanoma Than Number of Colors: A Possible Fractal-Based Explanation

Julio Urenda

The University of Texas at El Paso, jcurenda@utep.edu

Olga Kosheleva

The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich

The University of Texas at El Paso, vladik@utep.edu

Follow this and additional works at: https://scholarworks.utep.edu/cs_techrep



Part of the [Computer Sciences Commons](#)

Comments:

Technical Report: UTEP-CS-20-98

Recommended Citation

Urenda, Julio; Kosheleva, Olga; and Kreinovich, Vladik, "Why Number of Color Difference Works Better In Detecting Melanoma Than Number of Colors: A Possible Fractal-Based Explanation" (2020).

Departmental Technical Reports (CS). 1508.

https://scholarworks.utep.edu/cs_techrep/1508

This Article is brought to you for free and open access by the Computer Science at ScholarWorks@UTEP. It has been accepted for inclusion in Departmental Technical Reports (CS) by an authorized administrator of ScholarWorks@UTEP. For more information, please contact lweber@utep.edu.

Why Number of Color Difference Works Better In Detecting Melanoma Than Number of Colors: A Possible Fractal-Based Explanation

Julio Urenda^{1,2}, Olga Kosheleva³, and Vladik Kreinovich²

¹Department of Mathematical Sciences

²Department of Computer Science

³Department of Teacher Education

University of Texas at El Paso

500 W. University

El Paso, TX 79968, USA

jcurenda@utep.edu, olgak@utep.edu, vladik@utep.edu

Abstract

At present, the best way to detect melanoma based on an image of a skin spot is to count the number of different colors in this image. A recent paper has shown that the detection can improve if instead of the number of colors, we use the difference between numbers of colors computed by using different thresholds. In this paper, we provide a possible fractal-based explanation for this empirical fact.

1 Formulation of the Problem

Melanoma is a serious problem. Melanoma – one of the most frequent types of skin cancer – is ubiquitous in sun-exposed areas like Texas;

- when detected early, it is curable, but
- if it is not detected early, it can metastasize and kill the patient.

One way to detect early melanomas – and to distinguish them from harmless moles – is to analyze the image of the corresponding skin area.

How melanoma is detected now based on an image of a skin spot. One of the main techniques for detecting melanoma in an image is counting number of colors.

Specifically, there is a notion of a distance in the space of all possible colors. To count the number of colors, we select a threshold value $\varepsilon > 0$. We say that two colors are ε -*distant* if the distance between them in the color space is at

least ε . By the *number of colors*, we then mean the largest possible number of points for which every two points have ε -distant colors.

The larger the number of colors, the higher the probability that the patient had a melanoma. This method is good, but not perfect: there are:

- false negative cases – when melanoma has a small value of $N(\varepsilon)$, and
- false positive cases, when a benign mole has a large value of $N(\varepsilon)$.

Number of color difference: a new technique. A recent paper [1] showed that we can make this detection more efficient – with fewer false negatives and fewer false positives – if:

- instead of considering the number of colors $N(\varepsilon)$,
- we consider a new characteristic called the *number of color difference*: namely, the difference $N(\varepsilon_1) - N(\varepsilon_2)$ between numbers of colors corresponding to two different thresholds ε_i .

Formulation of the problem. The new technique is more efficient, but why?

What we do in this paper. In this paper, we provide a possible fractal-based explanation for the empirical success of the new characteristic.

2 Main Idea and the Resulting Explanation

Informal idea. There are several informal criteria for separating melanoma from benign skin formations. Some of them are based on dynamics, but some are based on the current state of the formation, namely:

- the larger the spot and
- more irregular (non-smooth) the spot,

the higher the probability that this spot is malignant. Similar idea works in a color space:

- the larger the set of points in the color space, and
- the more irregular this set,

the higher the probability that the spot is malignant.

What does irregular mean? The intuitive notions of smoothness vs. non-smoothness can be captured by the notion of *fractal dimension*, i.e., a Hausdorff dimension which is not an integer; see, e.g., [2]. Let us briefly recall this notion.

For smooth lines and surfaces, dimension can be defined as follows. For points on a smooth curve, asymptotically, for small ε , we have $N(\varepsilon) \sim A \cdot \varepsilon^{-1}$. For example, for an interval $[a, b]$, the largest number of ε -distant points can be obtained if we place:

- the first point at a ,
- the second at exactly the distance ε from a , i.e., at the point $a + \varepsilon$,
- the third at the point $a + 2\varepsilon$, etc.

This way, we can fit $\approx (b - a)/\varepsilon$ points.

Similarly, for a smooth surface, we have $N(\varepsilon) \sim A \cdot \varepsilon^{-2}$. For a region in the 3-D space, we have $N(\varepsilon) \sim A \cdot \varepsilon^{-3}$. In all these cases, we have $N(\varepsilon) \sim A \cdot \varepsilon^{-\alpha}$, where α is the dimension of the corresponding region.

It turns out that for irregular curves and surfaces, we often have $N(\varepsilon) \sim A \cdot \varepsilon^{-\alpha}$ for non-integer α . Since in the smooth case α is equal to dimension, the value α corresponding to non-smooth lines and surfaces is also called dimension. To distinguish it from other definitions of dimension, this value is known as *Hausdorff dimension*.

For example, for the trajectory of a Brownian motion, we have $N(\varepsilon) \sim A \cdot \varepsilon^{-1.5}$, so its Hausdorff dimension is 1.5.

Resulting idea. We assume that $N(\varepsilon) \sim A \cdot \varepsilon^{-\alpha}$, and we try to extract the value A and α from the observations.

To be able to better distinguish between benign and malignant formations, we need to better distinguish between different values of A and α . In other words, we need to select a characteristic C which is the most sensitive to changes in A and α , i.e., for which the relative changes

$$\delta_C(A) \stackrel{\text{def}}{=} \frac{\left| \frac{\partial C}{\partial A} \right|}{|C|} \quad \text{and} \quad \delta_C(\alpha) \stackrel{\text{def}}{=} \frac{\left| \frac{\partial C}{\partial \alpha} \right|}{|C|}$$

in C caused by changes in, respectively, A and α are the largest.

Let us compare the two characteristics from this viewpoint. The original characteristic is $N = A \cdot \varepsilon^{-\alpha}$. The new characteristic is the difference between the values $N(\varepsilon)$ corresponding to two close values of ε . In general, for close values ε_i , we have $f(\varepsilon_1) - f(\varepsilon_2) \approx f'(\varepsilon) \cdot \Delta\varepsilon$, where $f'(\varepsilon)$ denotes the derivative and $\Delta\varepsilon \stackrel{\text{def}}{=} \varepsilon_1 - \varepsilon_2$. In particular, for our case, when $N(\varepsilon) = A \cdot \varepsilon^{-\alpha}$ and thus, $N'(\varepsilon) = -\alpha \cdot A \cdot \varepsilon^{-(\alpha+1)}$, we have

$$D \approx -\alpha \cdot A \cdot \varepsilon^{-(\alpha+1)} \cdot \Delta\varepsilon. \quad (1)$$

For the original characteristic N , we have $\frac{\partial N}{\partial A} = \varepsilon^{-\alpha}$, so

$$\delta_N(A) = \frac{\left| \frac{\partial N}{\partial A} \right|}{|N|} = \frac{|\varepsilon^{-\alpha}|}{|A \cdot \varepsilon^{-\alpha}|} = \frac{1}{A}. \quad (2)$$

We also have

$$\frac{\partial N}{\partial \alpha} = -A \cdot \varepsilon^{-\alpha} \cdot \ln(\varepsilon),$$

so

$$\delta_N(\alpha) = \frac{\left| \frac{\partial N}{\partial \alpha} \right|}{|N|} = \frac{|-A \cdot \varepsilon^{-\alpha} \cdot \ln(\varepsilon)|}{|A \cdot \varepsilon^{-\alpha}|} = |\ln(\varepsilon)|. \quad (3)$$

Since usually $\varepsilon < 1$, we have $\ln(\varepsilon) < 0$, hence

$$\delta_N(\alpha) = |\ln(\varepsilon)| = -\ln(\varepsilon). \quad (4)$$

For the new characteristic (1), we have $\frac{\partial D}{\partial A} = \alpha \cdot \varepsilon^{-(\alpha+1)} \cdot \Delta\varepsilon$, so

$$\delta_D(A) = \frac{\left| \frac{\partial D}{\partial A} \right|}{|D|} = \frac{|-\alpha \cdot \varepsilon^{-(\alpha+1)} \cdot \Delta\varepsilon|}{|-\alpha \cdot A \cdot \varepsilon^{-(\alpha+1)} \cdot \Delta\varepsilon|} = \frac{1}{A}. \quad (5)$$

We also have

$$\frac{\partial D}{\partial \alpha} = [-A \cdot \varepsilon^{-(\alpha+1)} + \alpha \cdot A \cdot \varepsilon^{-(\alpha+1)} \cdot \ln(\varepsilon)] \cdot \Delta\varepsilon,$$

so

$$\delta_D(\alpha) = \frac{\left| \frac{\partial D}{\partial \alpha} \right|}{|D|} = \frac{|[-A \cdot \varepsilon^{-(\alpha+1)} + \alpha \cdot A \cdot \varepsilon^{-(\alpha+1)} \cdot \ln(\varepsilon)] \cdot \Delta\varepsilon|}{|-\alpha \cdot A \cdot \varepsilon^{-(\alpha+1)} \cdot \Delta\varepsilon|} = \left| \frac{1}{\alpha} - \ln(\varepsilon) \right|. \quad (6)$$

Here, $\alpha > 0$ and $\ln(\varepsilon) < 0$, so

$$\delta_D(\alpha) = \frac{1}{\alpha} + |\ln(\varepsilon)|. \quad (7)$$

Conclusions. By comparing sensitivities with respect to parameter A , i.e., formulas (2) and (5), we conclude that the old and the new characteristics have the exact same sensitivity – so in this sense, the new characteristic is not worse than the old one.

On the other hand, by comparing sensitivities with respect to parameter α , i.e., formulas (4) and (7), we see that the new characteristic is clearly more sensitive than the old one – which explains why the new characteristic is empirically better.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), and HRD-1834620 and HRD-2034030 (CAHSI Includes).

References

- [1] H. Akan and M. Z. Yildiz, “Development of new descriptor for melanoma detection on dermoscopic images”, *Medical and Biological Engineering and Computing*, 2020, PMID: 32865764.
- [2] B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman Publ., San Francisco, California, 1983.