

Automated detection of connective tissue by tissue counter analysis and classification and regression trees

Josef Smolle and Peter Kahofer

Department of Dermatology, University of Graz,
Auenbruggerplatz 8, A-8036 Graz, Austria
Tel.: +43 316 385/2530; Fax: +43 316 385/2466;
E-mail: josef.smolle@kfunigraz.ac.at

Received May 2001

Accepted 15 February 2002

Objective: To evaluate the feasibility of the CART (Classification and Regression Tree) procedure for the recognition of microscopic structures in tissue counter analysis.

Methods: Digital microscopic images of H&E stained slides of normal human skin and of primary malignant melanoma were overlaid with regularly distributed square measuring masks (elements) and grey value, texture and colour features within each mask were recorded. In the learning set, elements were interactively labeled as representing either connective tissue of the reticular dermis, other tissue components or background. Subsequently, CART models were based on these data sets.

Results: Implementation of the CART classification rules into the image analysis program showed that in an independent test set 94.1% of elements classified as connective tissue of the reticular dermis were correctly labeled. Automated measurements of the total amount of tissue and of the amount of connective tissue within a slide showed high reproducibility ($r = 0.97$ and $r = 0.94$, respectively; $p < 0.001$).

Conclusions: CART procedure in tissue counter analysis yields simple and reproducible classification rules for tissue elements.

Keywords: Image analysis, malignant melanoma, connective tissue, classification and regression tree, tissue counter analysis

1. Introduction

Though automated image analysis works well in cytology with isolated cells which can be more or less easily discriminated [7,15], automated evaluation of

histological slides is often hampered by the complexity of the microscopic scenes. Since completely computerized discrimination of certain structures often does not yield satisfactory results, extensive user interaction is necessary or a manual approach is preferred [4,8].

Previously we have suggested to avoid a priori discrimination and to divide digital microscopic images into elements of equal size and shape, followed by measurement of a set of image analysis features for each of the elements [12,13]. In a learning process, the elements are labeled as belonging to a particular tissue class, and statistical procedures are subsequently used to characterize the features characteristic for a particular class of elements. Based on this learning data set, algorithms can be derived which facilitate automated classification of new elements. This process has been termed "tissue counter analysis", since it "counts" the number of elements belonging to a particular class of tissue elements [12,13]. This approach does not only depend on the values of a single pixel, but also on a pre-defined neighborhood of that pixel, as it has been used in other studies [2,9]. The classification is then given to the whole neighborhood (element) and is not only assigned to the central pixel.

In previous studies, multivariate linear discriminant analysis has been used as a statistical classification tool. As a potential alternative, here we test the feasibility of CART (Classification and Regression Tree) analysis [5,14]. CART analysis has been developed to evaluate large data sets irrespective of the number of different clusters present within each class and has been applied to a wide variety of problems in natural and social sciences. In contrast to linear discriminant analysis, it does not depend on a certain distribution of the data and can lead to rather simple binary classification rules. In brief, CART procedures try to separate data sets into subsets termed "nodes" which show an optimal discrimination between the different classes that should be separated. Separation is performed by

searching for the best binary split criterion available. Each node is again tested for further splitting in order to gain a more accurate classification. When no further split is possible, the particular node is called a “terminal node” [5,14].

2. Material and methods

2.1. Specimens

A total of 36 skin biopsies was evaluated. 18 biopsies represented clinically normal appearing human skin taken from non-acral, non-facial skin. 18 biopsies represented primary cutaneous malignant melanoma. From each specimen, slides stained with haematoxylin and eosin were used. The largest section (in case of normal skin) and the section showing the largest vertical tumor diameter (in case of malignant melanoma) were selected and used for further analysis. 8 specimens each of normal skin and melanoma served as a learning set, and 10 specimens each were used as an independent test set.

2.2. Image analysis procedure

Analysis was performed using a Zeiss Axioskop bright field microscope (Zeiss, Oberkochen, Germany), mounted with a 3-chip colour video camera (Sony, Tokyo, Japan) connected to a KS 400 3.0 image analysis unit (Zeiss Vision, Hallbergmoos, Germany). Specimens were examined with a 20 \times objective resulting in a magnification of 0.66 μm per pixel. Each slide was scanned using a scanning table and an automated focusing device (Zeiss Vision).

For all measurements, illumination was kept constant at a background grey level of 200 (tolerance limit ± 4). Each measuring field was submitted to additive shading correction. No further image enhancement steps were performed.

An elementary measuring mask was defined as a square of 32 \times 32 pixels (21.12 \times 21.12 μm^2). In each 512 \times 512 pixel image, 64 elements at a square distance of 32 pixels were measured. For each element, mean, standard deviation, skewness, kurtosis, minimum and maximum were calculated in the intensity image and in the red, green and blue images. Furthermore, grey level moments and Haralick texture parameters were assessed [1], resulting in a total of 70 measuring features (Table 1).

In each slide, a measuring meander was interactively defined, which covered the whole epidermis, dermis, and part of the subcutis as well as the whole tumor area in cases of malignant melanoma.

Table 1

Measurement features obtained for each element. For detailed definition see [1]

Grey level features	Haralick texture features	Colour features
MEAND/grey	HARAM1	MEANDR/grey
STDD/grey	HARAM2	MEANDG/grey
SKEWD	HARAM3	MEANDB/grey
KURTD	HARAM4	STDDR/grey
ENERGYD	HARAM5	STDDG/grey
ENTROPYD/bit	HARAM6	STddb/grey
MIND/grey	HARAM7	SKEWR
MAXD/grey	HARAM8	SKEWG
SUMD/grey	HARAM9	SKEWB
SUMQD/grey ²	HARAM10	KURTR
MOMENTDC20/pixel ²	HARAM11	KURTG
MOMENTDC11/pixel ²	HARAR1	KURTB
MOMENTDC02/pixel ²	HARAR2	ENERGYR
MOMENTDC30/pixel ³	HARAR3	ENERGYG
MOMENTDC21/pixel ³	HARAR4	ENERGYB
MOMENTDC12/pixel ³	HARAR5	ENTROPYR
MOMENTDC03/pixel ³	HARAR6	ENTROPYG
MOMENTDI1	HARAR7	ENTROPYB
	HARAR8	MINR
	HARAR9	MING
	HARAR10	MINB
	HARAR11	MAXR
		MAXG
		MAXB
		SUMR
		SUMG
		SUMB
		SUMQR
		SUMQG
		SUMQB

2.3. Data set and CART analysis

In 8 specimens each of normal skin and melanoma, respectively, 10 randomly selected visual fields were scanned and each element was interactively labeled as belonging to background (area of the slide not covered by the section), connective tissue of the reticular dermis, or other tissue component. The data derived from all 16 specimens were pooled to form a learning set consisting of 10 240 elements. The data were stored using SPSS for Windows (SPSS Inc., Sunnyvale, USA) and transferred to the CART program package (Salford Systems, San Diego, USA). For this classification procedure, each of the 10240 tissue elements represents a “case” from a statistical point of view. In order to produce small and efficient trees, the smallest

node suitable for further splitting was defined as containing at least 200 elements, and the smallest value for a terminal node was set to at least 20 elements. Automated random splitting into learning and test subsets and tree optimization by evaluating 10 different provisional trees were used to create unbiased classification trees and to avoid splits which might occur by chance in random data. For this purpose, the original learning data are randomly split into 10 subsets, and each subset is used to create a tree. Only splits that turned out to be reproducible in these subsets are included into the final tree.

Two tree models were created: The first model was aimed to discriminate background elements from all tissue elements, and the second model to discriminate connective tissue of the reticular dermis from other tissue elements.

2.4. Automated application

The classification rules derived from the CART trees were implemented into the image analysis procedure as measurement conditions: Only elements fulfilling criteria specified by a particular sequence of rules – thus defined as belonging to a particular node – were measured. In 10 specimens each of normal skin and melanoma, 100 to 200 visual fields regularly distributed across the section were automatically evaluated and the elements belonging either to connective tissue of the reticular dermis or to other tissue components were stored in different data bases. For visual control of the classification procedure, the boundaries of the identified elements were displayed in colour in the overlay. The area occupied by all tissue elements and the area covered by connective tissue elements of the reticular dermis was calculated by multiplying the number of elements encountered with the area of the square element and a factor taking into account the spacing of elements and visual fields. Results are given in mm^2 .

2.5. Statistical methods

Mean value, standard deviation, standard error of the mean and linear regression analysis were used where appropriate [11]. Furthermore, the results of CART analysis were compared with multivariate linear discriminant analysis. These statistical evaluations were performed using the SPSS for Windows statistical program package (SPSS Inc.).

3. Results

3.1. General observations

Connective tissue of the reticular dermis was defined as coarse collagen bundles with occasional nuclei of fibrocytes. Blood vessels, nerves and arrector pili muscles embedded in the reticular dermis were labeled as other tissue components, as well as all other structures of the skin (epidermis, adnexal epithelium, adventitial dermis). Elements containing more than one component were labeled according to the major component present.

3.2. Background and connective tissue classification

In the CART tree discriminating background from tissue, 97.9% of background and 97.1% of tissue elements were correctly classified. In the CART tree discriminating connective tissue from other tissue components, 94.3% of connective tissue and 89.7% of other tissue elements were correctly classified. In both trees the majority (more than 90%) of the particular elements were found in a single terminal node (Fig. 1). The splitting rules defining these nodes were subsequently used for automated classification in the image analysis procedure. Background elements were characterized by two splits ($\text{ENTROPYB} = \text{entropy of the blue channel} > 4.205$ AND $\text{MINB} = \text{minimum of the blue channel} > 5.5$), while connective tissue of reticular dermis was characterized by three splits ($\text{ENERGYG} = \text{energy of the green channel} \leq 0.035$ AND $\text{MEANDR} = \text{mean of the red channel} > 153.48$ AND $\text{SKEWB} = \text{skewness of the blue channel} > -0.645$) [10].

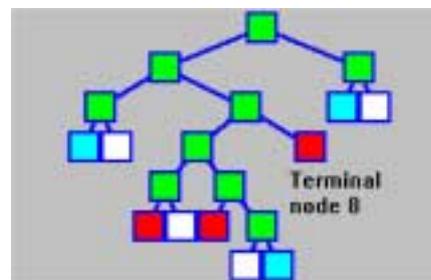


Fig. 1. CART (Classification and Regression Tree) for the detection of connective tissue of the reticular dermis, derived from the learning set of 10 240 elements. Non-terminal nodes appear green, terminal nodes consisting mainly of connective tissue elements appear red, terminal nodes consisting of other tissue elements appear blue. Nodes which do not contain a significant majority of one of the classes appear pale. The majority of elements classified as connective tissue is found within one terminal node (terminal node 8).

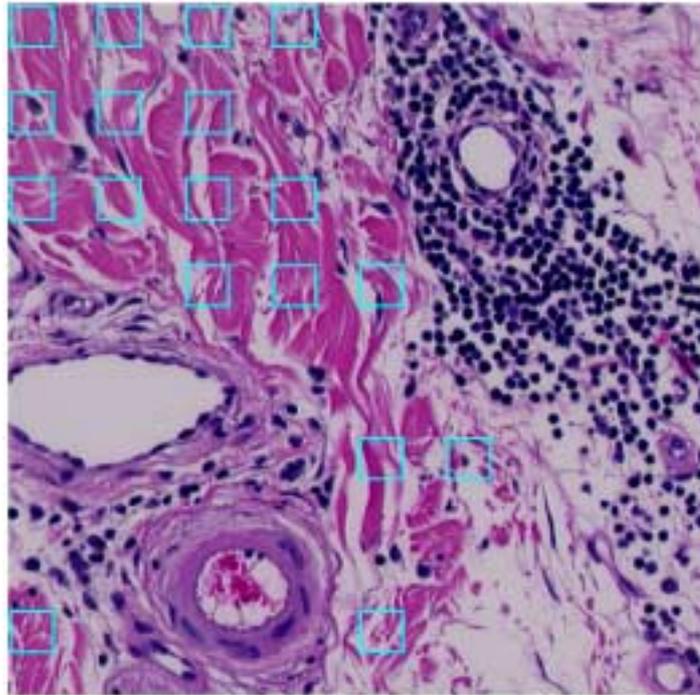


Fig. 2. Tissue counter analysis based on CART evaluation. Elements classified as connective tissue of the reticular dermis are indicated as squares. Loose adventitial connective tissue, cell-rich areas and inflammatory infiltrates as well as vessel walls are automatically excluded.

In five cases each of normal human skin and primary melanoma, the results of the classification procedure were checked by overlaying the elements recognized as dermal collagen tissue on the original image (Fig. 2) and by interactive counting of misclassified elements by an observer. It turned out, that $94.07 \pm 1.43\%$ were correctly classified. Elements erroneously classified as dermal collagen mainly represented horny layer, mature sebocytes, occasionally arrector pili muscle and accumulations of erythrocytes.

For comparison, multivariate linear discriminant analysis was also performed. It yielded a correct classification in 95.3% of background elements, 93.0% of connective tissue elements of the reticular dermis and 87.3% of other tissue elements.

3.3. Automated measurements

All 20 cases of the test set were evaluated automatically, with 100 to 200 visual fields distributed regularly across the section. User interaction was limited to setting the corners of the measuring meander and adjusting illumination to the desired value. The measuring procedure for one slide took about 5 min. Total amount of tissue elements was $9.8 \pm 1.4 \text{ mm}^2$ for nor-

mal skin and $93.8 \pm 21.7 \text{ mm}^2$ for primary melanoma, with the amount of connective tissue of the reticular dermis yielding values of $5.8 \pm 0.8 \text{ mm}^2$ and $28.5 \pm 5.3 \text{ mm}^2$, respectively. All measurements were carried out twice at different occasions. Correlation between consecutive measurements was $r = 0.97$ for the total amount of tissue and $r = 0.94$ for the amount of connective tissue (linear regression analysis: $p < 0.001$).

4. Discussion

The present study shows that tissue counter analysis together with CART (Classification and Regression Tree) evaluation can be applied to histologic sections and is able to detect certain tissue components. In contrast to multivariate linear discriminant analysis [6], the CART procedure leads to simple binary classification criteria which can easily be implemented in conventional image analysis programs.

Unlike to conventional approaches to image discrimination in histological slides, the relevant features need not be determined beforehand but are automatically detected by the statistical analysis procedure. As a prerequisite, however, a user-defined learning set is necessary.

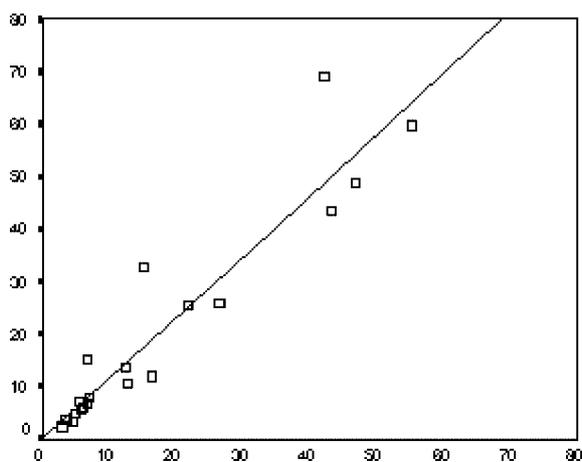


Fig. 3. Correlation between consecutive measurements – total area of connective tissue of the reticular dermis present within one section given in mm^2 as obtained in the test set of 10 cases each of normal human skin and primary melanoma (linear regression analysis: $r = 0.94$, $p < 0.001$).

The analysis of connective tissue of the reticular dermis as shown in this study may serve as an example for the application of tissue counter and CART analysis. As a main advantage, a fully automated procedure can be developed, with user interaction limited to setting the measuring meander. The results obtained by the proposed procedure yield highly reproducible results.

All measurement features of elements detected and classified by the measuring process can be stored in data bases and can subsequently be used for further statistical analysis. For example, connective tissue elements derived from different body sites or different pathologic processes, or before and after various treatment modalities, can be compared with each other and probably diagnostically relevant features may be extracted. Furthermore, the spatial coordinates of each element can be stored and the element can be re-localized in the section, facilitating a visual interpretation of mathematically developed classifications.

Besides CART analysis, multivariate linear discriminant analysis was also performed for comparison. It turned out that the results of the latter method were slightly less reliable than that obtained by CART analysis. There is, however, another advantage of the CART procedure compared to multivariate linear discriminant analysis. CART provides simple binary classification criteria suitable for automated image analysis procedure, while the multivariate discriminant approach results in one or more discriminant equations with a large number of parameters and coefficients. Furthermore, CART is able to provide good classifica-

tions also in settings where one class is represented by two or more different clusters in the parameter space – a situation which often hampers the performance of multivariate linear discriminant analysis.

In the presented procedure, classification was only performed for square image elements of 32×32 pixels, and the elements were 32 pixels apart from each other both in x and y direction. Theoretically, the measuring mask could be shifted in single pixel steps resulting in about 4096 times more measurements. This would be extremely time consuming without a significant effect on the results. Similar as in point counting methods, the standard error of a proportion p depends on the number of sampling points N according to the formula $SE(p) = ((p * (1 - p))/N)^2$ [3]. In case of 100 visual fields with 64 elements each, a tissue component with a proportion of 50% would yield a standard error of 0.6%, which is far lower than the empirically found error rate of about 6% in our procedure. Thus an increase of the number of measuring elements would not increase the accuracy of the counts. The segmentation result, however, which is a blocky one, could be improved by choosing smaller distances between the measuring masks.

The square size of 32×32 pixels as a measuring mask has been arbitrarily chosen. The number of different sizes, shapes and magnifications is virtually unlimited. It might be possible to enhance the reliability of the method by changing these parameters in a large scale investigation.

There are, of course, numerous limitations in the presented study. In particular, the number of cases is small, and variations of the image analysis setting have not been evaluated. Furthermore, the effect of different staining procedures or the variability between different laboratories needs to be addressed.

In conclusion, CART analysis may be a useful tool for the evaluation of data generated by tissue counter analysis.

References

- [1] KS 400 Imaging System Release 3.0, Carl Zeiss Vision, Munich, 1997.
- [2] W. Abmayr, E. Mannweiler, D. Oesterle and E. Deml, Segmentation of scenes in tissue sections, *Analyt. Quant. Cytol. Histol.* **9** (1987), 190–196.
- [3] W.A. Aherne and M.S. Dunnill, *Morphometry*, Edward Arnold, London, 1982.

- [4] M. Binder, I. Dolezal, K. Wolff and H. Pehamberger, Stereologic estimation of volume-weighted mean nuclear volume as a predictor of prognosis in "thin" malignant melanoma, *Journal of Investigative Dermatology* **99** (1992), 180–183.
- [5] L. Breiman, J. Friedman, R. Olshen and C. Stone, *Classification and Regression Trees*, Wadsworth, Pacific Grove, 1984.
- [6] F. Brosius, *SPSS 8.0. Professionelle Statistik unter Windows*, MITP-Verlag, Bonn, 1998.
- [7] H. Harms, H.M. Aus, M. Haucke and U. Gunzer, Segmentation of stained blood cell images measured at high scanning density with high magnification and high numerical aperture optics, *Cytometry* **7** (1992), 522–531.
- [8] G. Leitinger, L. Cerroni, H.P. Soyer, J. Smolle and H. Kerl, Morphometric diagnosis of melanocytic skin tumors, *Am. J. Dermatopathol.* **12** (1990), 441–445.
- [9] J.P. Rigaut, Automated image segmentation by mathematical morphology and fractal geometry, *J. Microsc.* **150** (1988), 21–30.
- [10] J.C. Russ, *The Image Processing Handbook*, Vol. 2, CRC Press, Boca Raton, Ann Arbor, London, Tokyo, 1995.
- [11] L. Sachs, *Statistische Methoden*, Vol. 4, Springer, Berlin, Heidelberg, New York, 1979.
- [12] J. Smolle, Computer recognition of skin structures using discriminant and cluster analysis, *Skin Research and Technology* **6** (2000), 58–63.
- [13] J. Smolle, Diagnostic assessment of cutaneous melanoma and common nevi using tissue counter analysis, *Analyt. Quant. Cytol. Histol.* **22** (2000), 299–306.
- [14] D. Steinberg and P. Colla, *CART: Tree-structured non-parametric data analysis*, Salford Systems, San Diego, 1995.
- [15] W. Stolz, T. Vogt, M. Landthaler, S. Hempfer, P. Bingle and W. Abmayr, Differentiation between malignant melanomas and benign melanocytic nevi by computerized DNA cytometry of imprint specimens, *J. Cutan. Pathol.* **21** (1994), 7–15.



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

