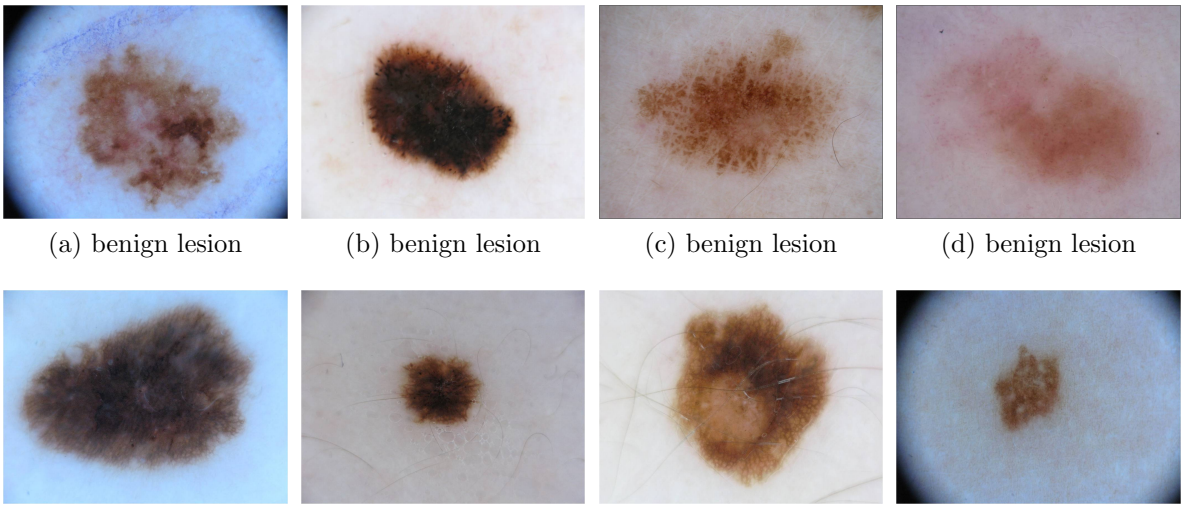
Project: Automatic classification of skin lesions (melanoma detection)

Note

The objective of this project is to propose a methodology for the automatic classification of skin lesions, based on image analysis and machine learning.

In 2012 there were more than 11150 new cases of skin cancer in France (3.1% of all detected cancers), 15% of which were mortal. Late diagnosis of skin cancer makes treatments much less efficient (i.e. a melanoma may become very aggressive in just a few months). Therefore, its early detection becomes essential to improve the chances of curing the skin cancer and, thus, the survival chances of the patient.

This is not an easy task for a non-experienced observer, as it is evidenced by the example shown in figure 2. To carry out the detection of melanoma, dermatologists use several state-of-the-art methods, often called *rules*, such as the ABCD rule [3] (Asymmetry, Border irregularity, Color irregularity and Differential structure, i.e. the size and number of structural features), the Menzies scoring or the 7-point checklist [1], which are based on the presence of certain texture patterns. Dermatologists detect malignancy features based on these rules and then combine this information to predict malignancy of the tumor.



(a) benign lesion (b) benign lesion (c) benign lesion (d) benign lesion

(e) melanoma (f) melanoma (g) melanoma (h) melanoma

Figure 1: Benign lesions vs. melanoma.

An automation of such individual analysis of nevi would be very helpful not only for dermatologists, as they would make more accurate diagnosis and, thus, make better decisions about the need of surgery, but also for general practitioners, sending patients to the specialist

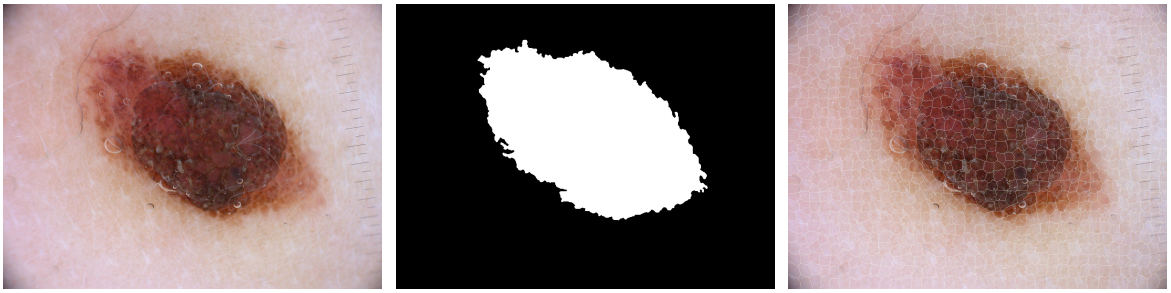
more appropriately. Therefore, it would save economic and time resources to the Social Security system.

Automatic melanoma detection – based on the analysis of dermoscopy images – has been receiving an increasing attention in the literature [2]. Most works reproduce the classic rules used by dermatologists, extracting features linked to them (see for instance [7]). However, the extraction of such kind of features needs a previous segmentation of the lesion, which is often considered a very tricky step, as it may not be clear, even for specialists, where the boundaries of the lesion are. Other methods have been investigated in the literature where dermoscopic images do not need any segmentation process as proposed in [8]. Such a method uses a texture descriptor computed with the Local Binary Patterns [5, 4, 6] (LBPs). Nethertheless, only intensity information of the color components are taken into consideration.

1 Database

The database is composed of 200 original images of pigmented skin lesions acquired by dermoscopy (selected from the ISIC 2017 database). 100 of them have been histopathologically confirmed as melanomas. Thus two classes are considered in this project: on the one hand confirmed melanomas and on the other hand the remaining benign lesions (look at the column denoted 'melanoma' of the file `ISIC_2017_Data_GroundTruth_Classification.csv`).

In addition to the original images, binary images representing the region of interest of each lesion are included in the database. The segmentation of the images using superpixels are also added. Pay attention to the number within the filename of the images and its associated class in the csv file (this latter contains the class of each image for the whole database ISIC 2017).



(a) original image (b) segmented lesion (c) superpixels

Figure 2: Segmentation of the skin lesion.

1. Load the different images from the database with suitable names.

2 Features extraction

In this part, the objective is to extract some features for each image of the database. You can use geometrical / morphological descriptors as well as intensity / texture descriptors.

1. For each image extract geometrical / morphological descriptors using the binary segmented image.
2. For each image extract intensity / texture descriptors using both the original and superpixel images.
3. Combine these descriptors.

3 Classification

Now you can use machine learning tools to classify the images in two classes: benign lesions or melanomas.

1. Divide your database in training and test sets.
2. Use classical machine learning tools (bayesian, SVM, neural networks...) to carry out the learning process using the extracted features.
3. Test you classifier on the test dataset.
4. Evaluate the performance of your classifier with quantitative criteria (confusion matrix, accuracy, F1 score, ROC curve...)

References

[1] R.H. Johr. Dermoscopy: alternative melanocytic algorithms—the ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist. *Clinics in Dermatology*, 20(3):240–247, may 2002.

[2] K. Korotkov and R. Garcia. Computerized analysis of pigmented skin lesions: A review. *Artificial Intelligence in Medicine*, 56(2):69 – 90, 2012.

[3] F. Nachbar, W. Stolz, T. Merkle, A.B. Cognetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, and G. Plewig. The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, 30(4):551–559, apr 1994.

[4] T. Ojala, M. Pietikäainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, Jul 2002.

[5] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51 – 59, 1996.

[6] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen. *Computer Vision Using Local Binary Patterns*. Springer, 2011.

[7] A. Tenenhaus, A. Nkengne, J.F. Horn, C. Serruys, A. Giron, and B. Fertil. Detection of melanoma from dermoscopic images of naevi acquired under uncontrolled conditions. *Skin research and technology*, 16(1):85–97, 2010.

[8] Y. Wazaefi, S. Paris, and B. Fertil. Contribution of a classifier of skin lesions to the dermatologist's decision. In *Image Processing Theory, Tools and Applications (IPTA), 2012 3rd IEEE International Conference on*, pages 207–211, Oct 2012.