

Project 2: Medical Imaging

Veronika Cheplygina, Amelia Jiménez-Sánchez

March 28, 2023

1 Overview

In this project you will learn to measure features in images of skin lesions, and predict the diagnosis (for example, melanoma) from these features in an automatic way. You will:

- Use Github to maintain your Python code
- Explore the dataset
- Implement methods to measure image features
- Explore and transform these features
- Use the features with machine learning classifiers to predict the lesion diagnosis
- Perform experiments to evaluate different parts of your method
- Write a report about your findings in LaTeX

2 Assignment

A dermatologist in your city asks if you can investigate whether some features of skin lesions can be reliably measured by image analysis algorithms.

Your goal is to do a small investigation, where you process the images to measure different types of features, train machine learning classifiers to predict the lesion diagnosis, and report on your results.

You must also deliver a trained classifier, that the dermatologist will evaluate on external images, that you do not have access to.

3 Project vs Exercises

This file contains some main tasks in your project. The tasks are a guideline and you do not have to do all suggestions in each task. If you do the minimum for each task, you can expect that the report has a minimum passing grade if everything is done correctly. You are therefore encouraged to go into depth with at least some of the tasks, exploring different options, integrating what you learn in the course, and using a bit of creativity.

The exercises (provided in separate files in some weeks of the project) do not always directly relate to the project, but can help you complete parts of it. You can do the exercises together with your project group if you want. There are no hand-ins for the exercises.

4 Task 0: explore the data

You will work with (part of) the public dataset PAD-UFES-20 [1], you can download the data here <https://data.mendeley.com/datasets/zr7vgbcyr2/1>.

Go through the data (images and meta-data) that you have available to understand what's available to you and write a brief summary of your findings. For example:

- What types of diagnoses are there, and how do they relate to each other? You can use Google Scholar or other sources, to learn more background about these diagnoses.
- Is there some missing data? Are there images of low quality? Etc.
- **You can select a small (e.g. 100) subset images to focus on at the start.** The only requirement is that there are at least two categories of images.

5 Task 1A: segment images

Create segmentations for some images. You can do this with image processing methods, or yourself with LabelStudio.

6 Task 1B: measure the features yourself

Search for related work about the Asymmetry and Color features and how they are measured by dermatologists.

Create an “annotation guide” for you and your group members, where you discuss at least 5 images together, and decide how to rate their Asymmetry and Color.

Then split the images, such that each image is annotated by at least two people in your group. Save your annotations in a CSV file, such that there are as many columns as there are different annotators (+ one column for the image name), i.e. do not put annotations of different people into the same column.

Make sure your CSV file follows the guidelines outlined in [?].

7 Task 2: measure the features with an algorithm

Create implementations for the Asymmetry and Color features using related work in image analysis. There will be multiple (similar) ways to measure each feature, if this is the case you can motivate which method you choose. You may use code available online but you need to be able to explain and modify different steps of the code.

To test your implementations, you might want to create “toy” images where you already know the results, for example a circle should be less asymmetric than an ellipse, etc.

Once you are satisfied with your implementations, run them on the real images and save the features in a CSV file.

Compare the features to your manual measurements by calculating agreement and/or visualizing the measurements. Do you agree with your algorithm? Do you see any other patterns?

8 Task 3: predict the diagnosis

For this task, you can use more images from the same dataset, or use other public data sources that you find.

Use a cross-validation setup to train different classifiers we studied in class (no neural networks), and evaluate their performance with appropriate metrics. You may also use other ways of evaluating classifiers, for example inspecting images that are classified incorrectly.

After this, select your best set of features + classifier. Train this classifier on the entire dataset (without cross-validation) and save the trained classifier.

Then create a function that can classify an external image/mask. This function should measure features you used, apply any transformations etc, and finally apply your trained classifier.

The classifier should output a probability of the image being suspicious, between 0 (healthy) and 1 (not healthy). This will be evaluated on a different set of data, which is not given to you. The external data will have masks available.

9 Task 4: open question

Use the data and your findings so far to formulate, motivate, answer, and discuss another research question of your choice. For example, you can study additional datasets, differences between groups of patients, additional types of features, etc.

10 Hand-in

You must hand in a report (PDF) and your Github repository.

10.1 Report

Write your report in LaTeX. You can include the following sections (variations on this are allowed):

- Add the link to your Github repository on the first page of the report
- Introduction: context and motivation for the problem. What has been already studied about this topic? What are your research questions?
- Data: describe your data, and any exploratory data analysis
- Methods: describe the methods you used for feature extraction, classification, and evaluation
- Results: what your results are, and possible reasons behind these results

- Discussion and conclusions: limitations of your approach, your conclusions about your research questions, general reflection (for example on ethics)

10.2 Github

The Github repository should contain:

- A .csv file with your manual ratings of the images
- A .csv file with your algorithm-measured features
- Code to reproduce the tables and figures in your report
- Your saved classifier
- A function that takes an image/mask and outputs its probability of being suspicious
- Any other code that you used

Do not include the PAD-UFES data in your repository.

References

- [1] Pacheco, A.G., Lima, G.R., Salomao, A.S., Krohling, B., Biral, I.P., de Angelo, G.G., Alves Jr, F.C., Esgario, J.G., Simora, A.C., Castro, P.B., et al.: Pad-ufes-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones. Data in brief **32**, 106221 (2020)