IT UNIVERSITY OF COPENHAGEN

First Year Project

Project #3: Image Analysis

Group 3:

Carl August Wismer (cwis@itu.dk)
Chrisanna Kate Cornish (ccor@itu.dk)
Danielle Marie Dequin (ddeq@itu.dk)
Gino Franco Fazzi (gifa@itu.dk)
Moneeca Abru Iftikhar Latif (abml@itu.dk)

2nd Semester, Spring 2021
Data Science

1 Introduction

Melanoma is a serious form of skin cancer, affecting over 350,000 people globally in 2015, resulting in almost 60,000 deaths (Wang et al.). It can be easy to treat and cure if detected early (Rigel and Carucci). However, melanomas can be difficult to distinguish from seborrheic keratosis, a common benign skin condition, which can lead to misdiagnoses and increased risk (Izikson et al.). As a response to this, many algorithms have been developed to detect if a lesion is a melanoma or seborrheic keratosis. These technologies are applied in smartphone applications ('apps'). According to one review, these apps have a high chance of giving false negatives, which could lead to a delay in appropriate treatment and thereby increasing risk (Chuchu et al.).

This report investigates the research question: "Is it possible to reliably determine if a lesion is melanoma, using the first three features of the ABC rule (Asymmetry, border irregularity and color)?" This research can be used to determine limitations in these technologies, which can be used to refine their development for more accurate and therefore safer application.

2 Data

2.1 Source

The International Skin Imaging Collaboration (ISIC) is an international effort to improve melanoma diagnosis (Codella et al.). The data comes from the 2017 dataset within the ISIC Archive and includes, per lesion, an image of the lesion, a mask of the image showing pixels that belong to the lesion, and a label that shows whether the lesion was diagnosed as melanoma or seborrheic keratosis. This archive contains the largest publicly available collection of quality controlled dermoscopic images of skin lesions. A total of 2750 observations were used for the research.

2.2 Data Cleaning

Visual inspection showed that some of the original images did not contain the entire lesion. As this study focused on lesion asymmetry and border irregularities, the complete border was required. Therefore, a function that went through the images was applied. If an image had a certain percentage of the lesion (20%) on the border, the corresponding image was selected for visual inspection. The images were then gone through manually, resulting in a total of 151 images (approximately 5,49% of the whole dataset) excluded from the research. Out of the 151 excluded images, 48 were melanoma, 24 were keratosis and 79 were neither melanoma or keratosis.

2.3 Image reduction

In order to save space and run a more efficient feature extraction, the color images were cropped using their segmentation mask and reduced to the minimum frame around the lesion. This reduced the data size, allowing for faster run time on the relevant functions, and also allowed for the calculation of the RGB channels on the lesion alone.

2.4 Feature extraction

An automatic feature extraction function was applied for the rest of the images:

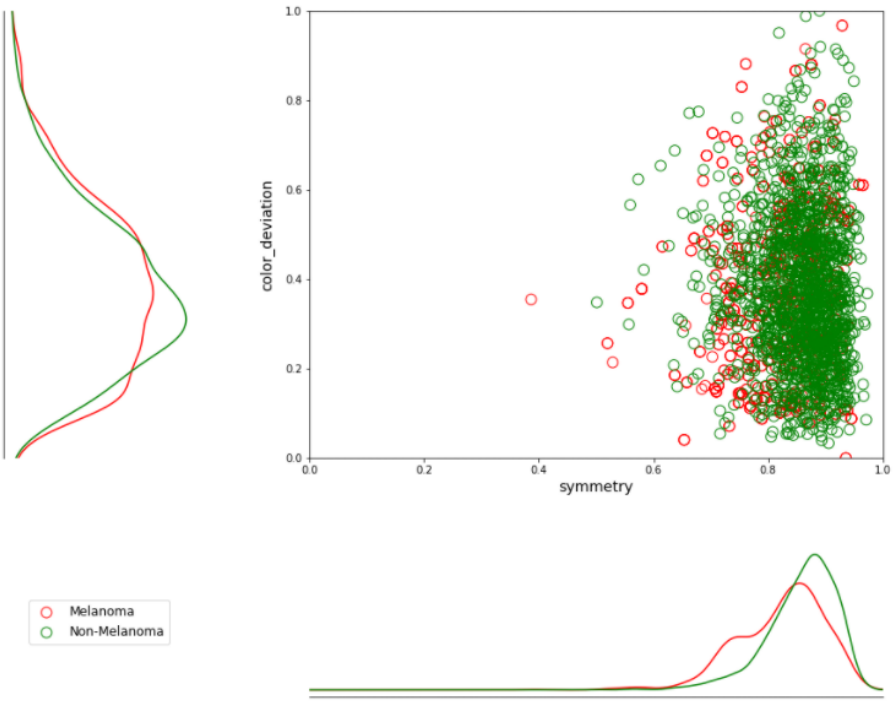


Figure 1: *Symmetry and Color*

- Symmetry: The images were rotated, cut vertically and horizontally, and superimposed; checking the top/bottom and left/right cuts for symmetry. The symmetry for both cuts was then averaged to obtain a unique metric. Symmetry ranges from 0 (non symmetric) to 1 (completely symmetric).
- Compactness: As a method to estimate the lesion border irregularity, a compactness method was applied. For this, perimeter and area is calculated, and compared to the theoretical relationship between these if the lesion was completely round (therefor without border irregularities). This metric returns a value between 0 (completely irregular) to 1 (completely regular).
- Color deviation: To estimate the color variation in the lesion, the standard deviation for each of the 3 RGB channels average was computed. As this metric returns the degree of variation for the RGB channel, it was not consistent with the range of the other features. The values were then scaled to match the 0 to 1 range with a Min-Max scaling formula:

$$scaled\, x = \frac{x - min(X)}{max(X) - min(X)}$$

3 Results

3.1 Plotting the Data

Scatter plots were created to visualize the patterns in the aggregated features data (*Figure 2*). As seborrheic keratoses are benign, we opt for a binary dissection of the observations: Melanoma or Non-Melanoma. Data were visualized using kernel density plots positioned alongside the scatter plots (*Figure 2*). There does not seem to be an obvious visual distinction between melanoma compared to non-melanoma in either the scatter plots or kernel density plots.

The data were then split into three sets; training set, validation set and test set. The training data consisted of 1932 total observations; 1578 non-melanoma and 354 melanoma (18%). To create a more balanced

dataset, the data were re-sampled using a replacement bootstrap method. This re-sampling provided 1578 non-melanoma and 966 (38%) melanoma observations to more accurately predict the results in data using the models that follow.

Two models were used for training and prediction; K-Nearest Neighbor and Decision Tree.

3.2 Model Training and Tuning

We will prioritize the recall score[1] for the specific "Melanoma" label, as we are primarily concerned on correctly identifying Melanomas.

3.2.1 K-Nearest Neighbors (KNN)

For the KNN model, the training data was fitted and tried on the validation data with a different number of neighbors until the optimal one was found (k=6). After training, this model was able to accurately predict 11 out of 24 melanomas in the validation set. The overall model accuracy is 60%, which did not instill much confidence in prediction.

- Recall: Our model achieved a recall score of 46%, detecting 11 out of the 24 TRUE Melanomas.
- Precision [2]: The KNN model revealed a 21% precision score for melanoma (11 TRUE positives over 52 predicted).

3.2.2 Decision Tree

The decision tree model was able to accurately predict 4 out of 24 melanomas in the validation data, with the following scores:

- Recall: The Decision Tree model achieved a recall score of 17%.
- Precision: The Decision Tree model revealed a 13% precision score for melanoma (4 TRUE positives over 30 predicted) in the validation set.

3.3 Model prediction on Test Data

After assessing our models with the validation data we decided to move forward with a K-Nearest Neighbors with a K of 6 predicting the TEST dataset, with an overall accuracy of 56.77%. Our final results were:

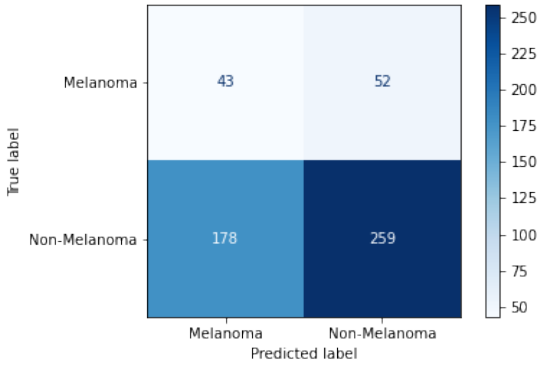
- Recall: The model achieved a recall score of 45%, predicting 43 of the 95 positive Melanomas.
- Precision: The model revealed a 19% precision score for melanoma (43 TRUE positives out of 221 predicted) in the test set.

[1] Recall: the ability of a classifier to find all positive instances.
[2] Precision: is the accuracy of positive predictions.

	precision	recall	f1-score	support
Melanoma	0.19	0.45	0.27	95
Non-Melanoma	0.83	0.59	0.69	437
accuracy			0.57	532
macro avg	0.51	0.52	0.48	532
weighted avg	0.72	0.57	0.62	532

Overall Model Accuracy: 56.77%

(a) KNN classification report on Test data



(b) Confusion matrix for KNN Model on Test data

Figure 2: *KNN Model results, K=6*

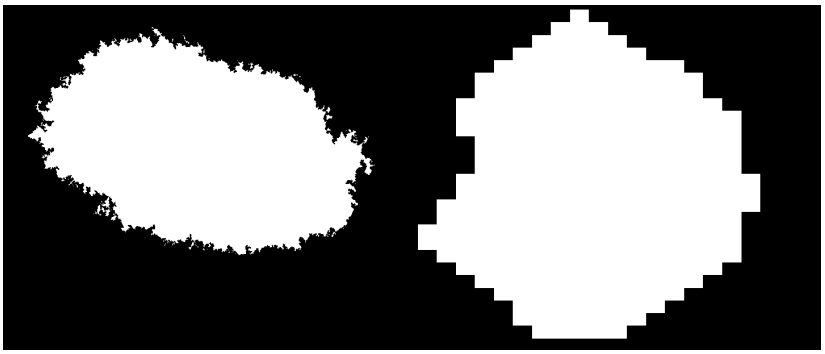


Figure 3: *Example of the huge variety in quality between two different segmented images*

4 Limitations

The data acquired for this research arrived to the ISIC Challenge from multiple sources, using non-uniform techniques. The images focus on different aspects of lesions, have varying quality, and are not all adequate for this analysis. For example, documenting using submersion in liquid allows at least one microstructure to be clearly visible for diagnosis (Chen et al.). Photographs made in this way can distort the image so that the border is not differentiable, leading to incorrect segmentation images and therefore inaccurate asymmetry and border analysis.

In addition, segmented images were produced with varying degrees of sensitivity, showing variation in quality and precision. Some segmented images are over simplified, while others have a significantly more precise border. This, too, would affect analysis of both asymmetry and border.

For future research of border and asymmetry, it would be ideal to have images made specifically for this analysis. For example, images should be made that show the entirety of the perimeter accurately. Furthermore, the techniques used to create segmentation images should be checked for accuracy before being submitted to an archive as a correct representative of the lesion. Given more time, it would have been beneficial for this research to develop and run testing using a more accurate process for segmentation.

While the algorithm developed for this research picked up on many of the images that were cut off, there were additional images that could have been filtered out before analysis due to inaccuracy in segmentation or poor image quality, as this can affect the accuracy and validity of the analysis performed.

While exploring the images, it was observed that the data do not include all skin tones. The present

techniques used to segment an image look for a strong change in color or tone to find the border of a lesion, and therefore would have to be tested for accuracy among all skin tones. It should be noted, however, that the type of melanoma that develops lesions discussed in this research is cutaneous melanoma. While this is the most common form, people who have darker skin tones are far less likely to develop this form, but are more likely to develop different types of melanoma that do not present as lesions (Melanoma Research Alliance).

Measuring symmetry with maximum precision is a very demanding task, especially with larger datasets. Images in our research were rotated 5 degrees at a time, to find the most symmetrical cut. The accuracy of the symmetry measurement could be improved by making more cuts, or by using other methods besides 5 degrees rotations.

5 Concluding remarks and recommendations

Although our analysis has shown very important limitations when using technology to efficiently detect and diagnose skin diseases, we cannot ignore that technology has a lot to offer to the medical world.

Many advantages may come from the use of image processing technologies as tools for helping both patients and doctors. Professionals could be assisted in image collection, process and labeling, enabling them to focus on other more important tasks that cannot be given to machines. On the other hand, if patients can have an efficient tool for keeping track of skin lesions, they could detect suspicious lesions earlier, improving the chances of a better long term outcome and more efficient treatment.

But in order to safely implement smartphone apps as an early self-detection tool for diseases like skin cancer, we must take a very thorough approach into the methodology and needs of the processed data. For this reason we reflected upon our own found limitations and drafted a non-exhaustive list of things to take into account:

- Photos should be taken within a specified range, keeping the entire lesion in frame, while still not zooming so far out that the lesion becomes small and pixelated.
- Photos should be taken with a clear source of light to allow a clear distinction of the colors of the lesion. If multiple photos need to be taken, a similar light source should be used for all the photos, as a light source with a different hue might change the photographed color of the lesion.
- Several photos should be taken within a time interval to adequately assess how the lesion is progressing.

5.1 Future Work

To develop this model further, more features could be included for example: age, sex and other potential risk factors. This could improve the accuracy of the app if the model is better able to identify which lesions are more likely to be melanoma.

Besides all these technical recommendations, people should be conscientious about the use and limitations of these apps, to safely convey the message that this is only a pre-diagnosis and should not replace a proper medical consultation.

References

Rigel, Darrell S, and John A Carucci. “Malignant melanoma: prevention, early detection, and treatment in the 21st century”. *CA: a cancer journal for clinicians* 50, no. 4 (2000): 215–236.

Izikson, Leonid, et al. "Prevalence of melanoma clinically resembling seborrheic keratosis: analysis of 9204 cases". *Archives of dermatology* 138, no. 12 (2002): 1562–1566.

Chen, Tzu-Hsiu, et al. "Comparison of visual effects of immersion fluids for dermoscopic examination of acral volar melanocytic lesions". *Dermatologica Sinica* 32, no. 2 (2014): 69–74.

Wang, Haidong, et al. "Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015". *The lancet* 388, no. 10053 (2016): 1459–1544.

Codella, N, et al. "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)" (2017).

Chuchu, Naomi, et al. "Smartphone applications for triaging adults with skin lesions that are suspicious for melanoma". *Cochrane Database of Systematic Reviews*, no. 12 (2018).

Melanoma Research Alliance. *Melanoma & Skin of Color*. https://www.curemelanoma.org/about-melanoma/people-of-color/. Accessed: 2021-04-18, 2020.