Contents

1 Project Description **1**

2 Theoretical Foundations and Methods **3**

3 Algorithms **4**

3.1 Working Conditions 4

3.2 Implementation and Comparison 5

3.2.1 Preprocessing 6

3.2.2 Segmentation 7

3.2.3 Postprocessing 8

3.2.4 Support vector machine 8

4 Evaluation **9**

Reference **12**

1 Project Description

Melanoma is a malignant tumor of melanocytes and it is the most dangerous type of skin cancer. In 2015, 3.1 million people had active disease, which resulted in 59,800 deaths [1][2]. "Melanomas typically occur in the skin, and about 25% of melanomas develop from moles" (Wild, 2014) [3]. Changes in moles from melanoma include increased size, irregular borders, color changes, and itchy or broken skin [4].
Early signs of melanoma are changes to the shape or color of existing moles or, in the case of nodular melanoma, the appearance of a new lump anywhere on the skin. Early signs of melanoma are summarized by the mnemonic "ABCDE" [5]:

- **A**symmetry: The Border of moles is asymmetry.
- **B**order: irregular with edges and corners.
- **C**olor: variegated, multiple colors in moles.
- **D**iameter: greater than 6 mm.
- **E**volving over time

Figure 1.1 shows the identification of normal and abnormal moles by ABCDE rule. If people can know whether they have melanoma at an early stage, they can be treated earlier.

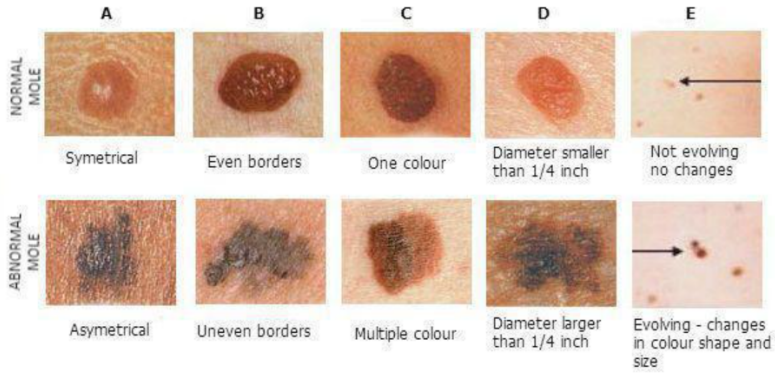


Figure 1.1: Difference Between Normal Moles and Abnormal Moles [6]

Our project is an automatic melanoma detector. Because the testing of this application requires expert knowledge, we were limited to working with a public dataset. We got

data from Skin Cancer MNIST: HAM10000 [7], which is the current state-of-the-art public resource for training in this domain. There are a total of 10015 images, and 961 of them are known to be melanomas. Because we cannot obtain comparative photos of the same mole changing in different periods, we use the **ABCD rule** in this project to detect whether a skin lesion is a melanoma.

Our program aims to serve two purposes. Firstly, to provide medical information to the general public for initial checks. Secondly, as a tool for experts to support their diagnostic decisions.

2 Theoretical Foundations and Methods

We'd like to develop a program that can automatically detect melanoma. If the user selects an image in the database, it will return a prediction value, whether the image is melanoma or non-melanoma.
In this work, we mainly searched the relevant literature on the automatic detection of melanoma [8][9][10], ABCD rules [11], and TDS [12].
These three articles [8][9][10] will be introduced more in detail in the following chapters.
The article "Automating the ABCD Rule for Melanoma Detection: A Survey" by Ali, Li, and Yang [11] summarized various methods for melanoma detection, this provided us with a lot of relevant methods for classifying and identifying melanoma.
The article "Fully Automated Approach for Early Detection of Pigmented Skin Lesion Diagnosis Using ABCD" by Mabrouk et al. [12] introduced in detail the method of detecting melanoma by Total Dermoscopic Score (TDS), also another way to detect melanomas.
The article "A Morphological Approach to Curvature-Based Evolution of Curves and Surfaces" by Pablo Márquez-Neila, Baumela, and Alvarez [13] which we used for image segmentation.
We also refer to other literature, such as "A computer-aided diagnosis system for malignant melanomas" by Razmjooy et al. [14]. This article introduced a very good way to remove noises cause by hairs. But after implementation, the result of the method is not ideal when dealing with lesions composed of multiple colors. If there is a very dark color, it can only extract the deepest area, not the entire lesion area where we are interested.
On the basis of each article, we consider the feasibility and expect better results. We tried different methods in segmentation and feature extraction, we will also show the results from different methods in chapter 3. At the same time, machine learning is used to classify in order to obtain higher accuracy results. In the end, users can select images from the dataset and get the prediction results.

3 Algorithms

In the preprocessing step, we introduce a Comparison between different methods for feature extraction and segmentation for diagnosing melanoma in order to find the methods with the highest accuracy. This could aid dermatologists in detecting melanoma skin cancer with higher accuracy since it is hard for the human eye to detect key features like asymmetry and color gradient.
SVM-based melanoma detection is the process of classifying 2 class problems, either benign or malignant Melanoma, hence even for an expert doctor some key features could be hard to detect for the naked eye.
In this chapter, we will first introduce the working conditions of our project, and then more details on implementation are described.

3.1 Working Conditions

Figure 3.1 shows the system flow of our project based on the article "Novel Approaches for Diagnosing Melanoma Skin Lesions Through Supervised and Deep Learning Algorithms" by Premaladha and Ravichandran. Based on the original, we have modified the flow chart to make it better adapt to our project.
The input of our program is colored images of a skin lesion with a fixed size of (600 × 450) pixels and metadata information for confirmation of the results. After segmentation and feature extraction, we used SVM to do the classification. The result of our program is a prediction of positive or negative skin lesions (Positive for melanoma and negative for not melanoma).

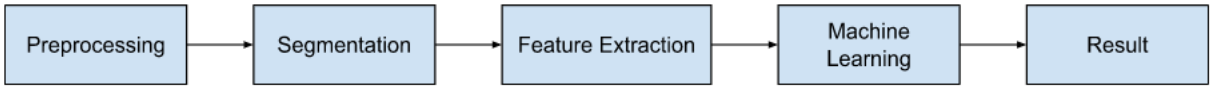


Figure 3.1: Pipeline of Melanoma Detection System [9]

There are several assumptions for our project:

- Every image has a closed contour that could be extracted, otherwise, it will be discarded from the detection method.
- Entire region of the interested lesion is visible.
- Only one class of melanoma is being diagnosed.

Unfortunately, some of the lesions that we are interested in are not fully visible in images. In this case, the result of segmentation is false and we can not classify this image.
In our work, there are 2 conditions for input images:

- The entire area of the skin lesion is visible and the color deviation is significant enough.
- There is no lens artifact.

Otherwise, in the preprocessing principle, the closed contour can not be detected or the lens will be detected as a part of a skin lesion. In the case of dark skin images also has limitations for image processing.
Figure 3.2 and Figure 3.3 show the false detection because the skin lesion in the input image is very disturbed or very bright, in both figures, the detector failed to lock on the lesion area due to high disturbance.

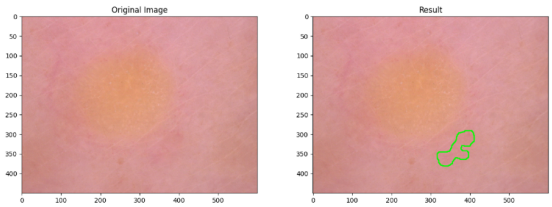


Figure 3.2: False Detection Example

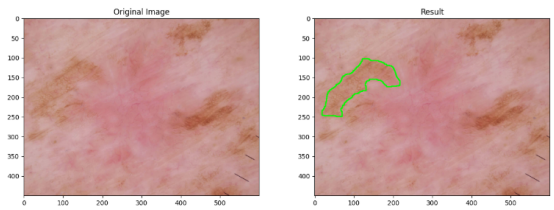


Figure 3.3: False Detection Example

Figure 3.4 shows false detection because the skin lesion is not completely visible in the image, the upper area of this lesion is missing. There are no features that can be extracted.

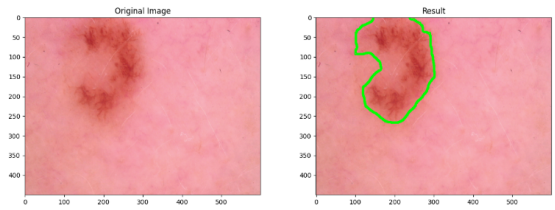


Figure 3.4: False Detection Example

3.2 Implementation and Comparison

Figure 3.5 shows the implementation steps of our project. Based on the system flow depicted in figure 3.1, the entire development project can be divided into four stages, and each step includes implementations:

1. Preprocessing: Gamma correction and blurring
2. Segmentation: Thresholding, Morphological transformations, and contours extraction
3. Postprocessing: Feature extraction
4. Support vector machine: Classification

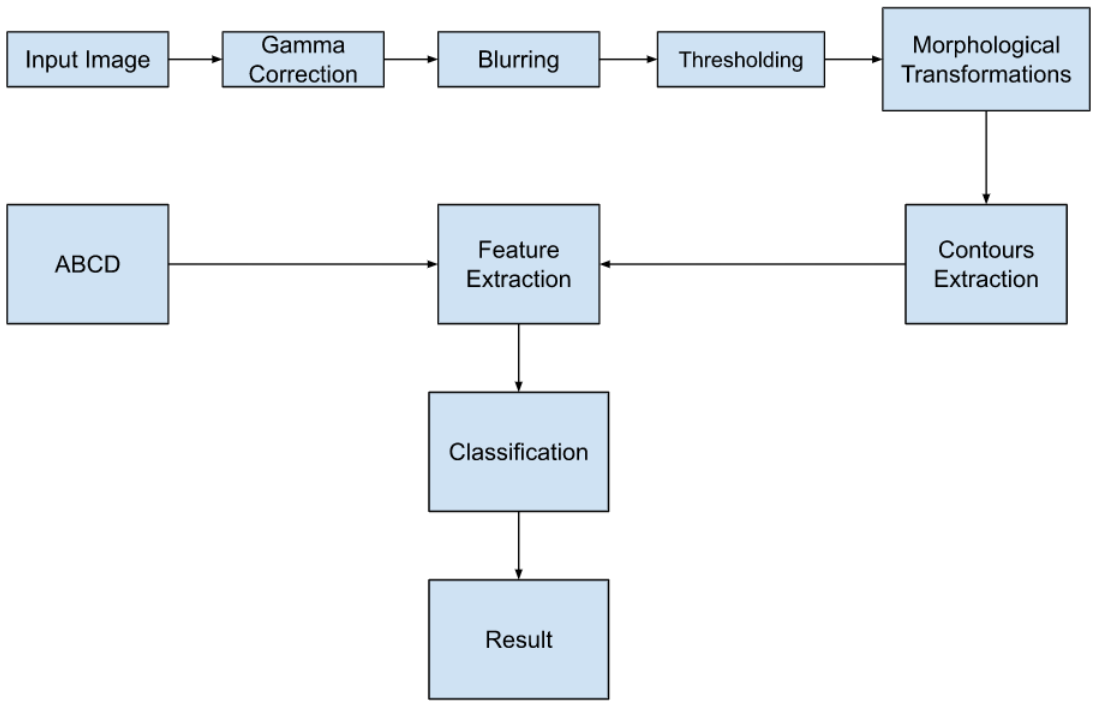


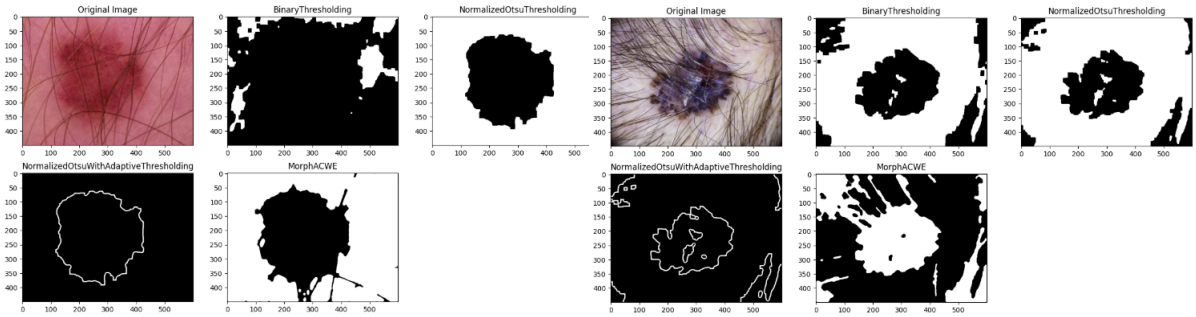
Figure 3.5: Implementation Steps of Project

3.2.1 Preprocessing

The images in our dataset already have the same size, we only need to adjust the brightness of the images in order to compensate for the non-uniform illumination in the image. We used gamma correction [8] to do this process.
For blurring a Bilateral Filter is used for preserving sharp edges, it uses two Gaussian functions, one for the coordinate space to make sure that only nearby pixels are considered for blurring while the other gaussian for color space ensures that only those pixels with similar intensities to the central pixels are considered for blurring, thus it preserves the edges since pixels at edges will have large intensity variation.

3.2.2 Segmentation

We used thresholding and morphological transformations in this step. In thresholding, 4 different approaches were tested and compared: Binary thresholding, normalized otsu thresholding, normalized otsu thresholding combined with adaptive thresholding, and morphological snakes active contours [13].
We wanted to remove noises (hairs in images) in the images, and only focus on the skin lesions for better results of feature extraction. In order to remove the noises caused by the hairs, we implemented two morphological operations: closing (dilation followed by erosion) and opening (erosion followed by dilation).
Figure 3.6 show the results using different thresholding techniques. Figure 3.7 shows the high efficiency of morphological operations for omitting the hairs' noises.



(a) (b)

Figure 3.6: Segmentations Using Different Thresholding

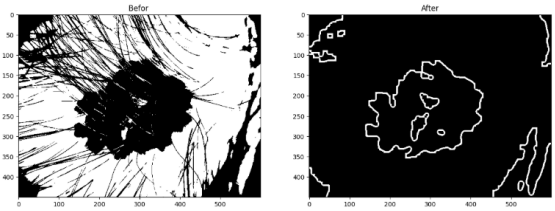


Figure 3.7: Remove hairs

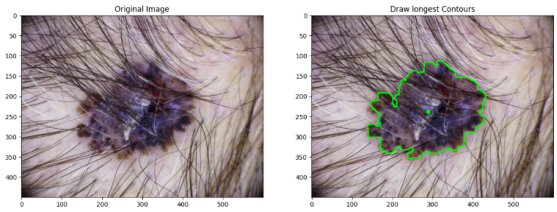


Figure 3.8: Find lesion's contour

In most cases, using binary Thresholding had the worst result, as in figure 3.6a. For better generalization, we found out that using 2-step thresholding: normalized otsu thresholding [9] combined with adaptive thresholding had the best results.
At the end of this step, we are only interested in the contours with the largest area. So we focus on that and extract the features for the SVM model.

3.2.3 Postprocessing

The algorithms of feature extraction are based on the article "Computer Aided Melanoma Skin Cancer Detection Using Image Processing" by Jain, Jagtap, and Pise [8] and article "Melanoma Skin Cancer Detection Using Image Processing and Machine Learning Techniques" by Thaajwer and Ishanka [10].
Based on [10], we extracted features using equation:

$$\begin{aligned} Asymmetry =& \frac{Perimeter}{Area} \\ Circularity =& \frac{Perimeter^2}{Area} \\ Abnormality =& \frac{4\pi \times Area}{Perimeter} \\ Irregularity =& \frac{Perimeter^2}{4\pi \times Area} \end{aligned} \tag{3.1}$$

Based on [8], we extracted four irregularity indexes as features by equations:

$$\begin{aligned} IrA =& \frac{Perimeter}{Area} \\ IrB =& \frac{Perimeter}{GD} \\ IrC =& Perimeter \times (\frac{1}{SD} - \frac{1}{GD}) \\ IrD =& GD - SD \end{aligned} \tag{3.2}$$

where GD is the greatest diameter and SD is the shortest diameter.
These features are used for training and classifying.

3.2.4 Support vector machine

The SVM principle are based on the article "Computer-Aided Diagnosis of Micro-Malignant Melanoma Lesions Applying Support Vector Machines" by Jaworek-Korjakowska [15] and the article "Novel Approaches for Diagnosing Melanoma Skin Lesions Through Supervised and Deep Learning Algorithms" by Premaladha and Ravichandran [9].
First approach was to use features from individual papers and later then we combined features from different papers and got a higher accuracy. We have tested independent features (for example Perimeter, ellipse irregularity, etc.) but that didn't lead to a better result.
As we mentioned in chapter 1, we have 10015 images in our dataset, and only 961 of them are melanomas. Only 9.6% of total images are positive class (class 1), and other images are negative class (class 0). Due to the high number of negative examples, the dataset was imbalanced. We had to select fewer negative samples so that the number of images of the positive class and negative class is almost the same, and then train the images.

4 Evaluation

The total time of training for all data is about 111 minutes, and the average time of each image is 668ms.

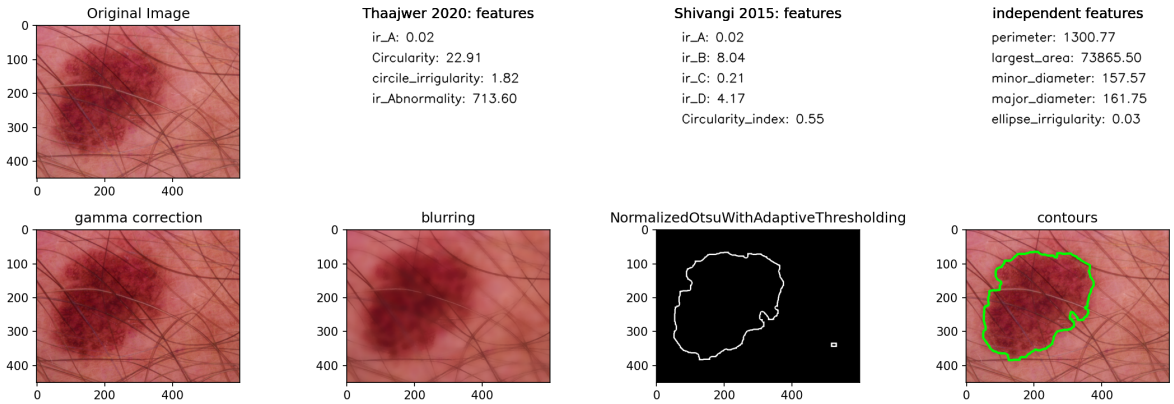


Figure 4.1: Result of Feature Extraction

Figure 4.1 shows the result of segmentation and feature extraction. We used normalized otsu thresholding combined with adaptive thresholding [9] and find the longest contour as the contour of a lesion, which is marked as green. The feature extraction result based on the article by Thaajwer and Ishanka [10] is shown as Thaajwer 2020 features. The result based on the article by Jain, Jagtap, and Pise [8] is shown as Shivangi 2015 feature. At last, the independent features are also shown in this figure.

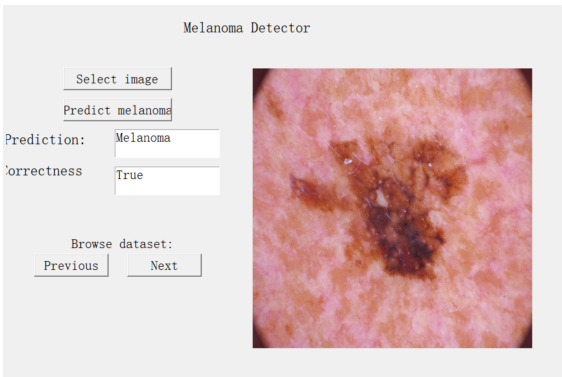
(a) (b)

Figure 4.2: Examples of Failed to Extract Features

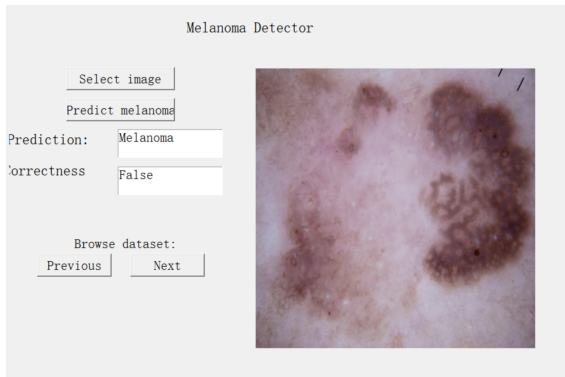
15.2% of total images failed to extract features because closed contours are not detected. Because a part of the lesion area is outside the images, there are no closed contours detected, see figure 4.2a. Or the picture has a lens effect so that the black areas at the four corners are considered to be part of the lesion, and there is also no closed contour detected, see figure 4.2b.

Our original idea was that when we select an image from the dataset, we get a result of either melanoma or not melanoma. Since in the dataset, we have already known whether the lesion in an image is melanoma, we also added the comparison between the predicted result and the real result at the end.

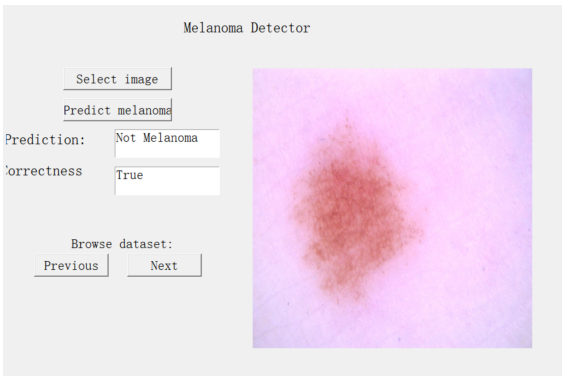
Figure 4.3 shows different results. Figure 4.3a and figure 4.3c are correctly detected. Figure 4.3b is not melanoma but detected as melanoma. The reason may be that after feature extraction, the border is asymmetrical and irregular. Similarly, the entire lesion area has different shades of color. It is consistent with the result of detecting whether it is melanoma based on the ABCD method. Figure 4.3d shows another example of the false result of the detection. Although this lesion is melanoma, its boundary is relatively regular and symmetrical, which affects the results of automatic detection.



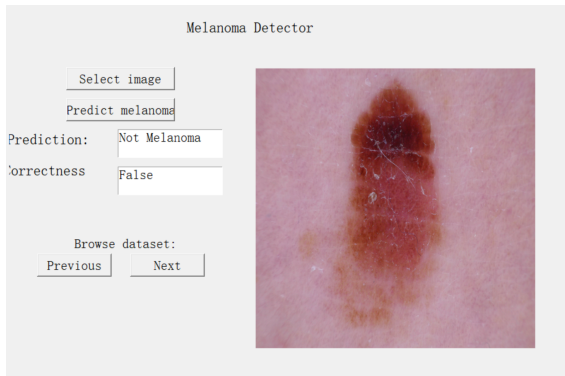
(a)



(b)



(c)



(d)

Figure 4.3: Four Different Results

In this project, we implemented an automatic melanoma detector. When we select an image from the dataset, as seen in figure 4.2 and figure 4.3, the program will return a result, whether the lesion is predictable, if predictable, what is the result and whether it is correct.
The probability of correct prediction is about 70%, which is obviously not a completely ideal result at the current stage. As we discussed before, such as figure 3.3 and figure 3.4, although the closed contour is detected, the contour area did not cover the entire lesion or the area of interest. This can lead to wrong results.
In future work, we hope to improve the accuracy of segmentation, Because obtaining contours with higher accuracy is a prerequisite for feature extraction. In the article "Fully Automated Approach for Early Detection of Pigmented Skin Lesion Diagnosis Using ABCD" by Mabrouk et al. [12] compared different results using different classification methods. We also want to try to use ways to do the classification and test the results.
We hope that after improving our project, people in the medical field can benefit from it, and we hope that individuals who have melanoma can perform self-examinations and seek medical treatment promptly.

Bibliography

[1] T, Vos ; C, Allen ; M, Arora ; RM, Barber ; ZA, Bhutta ; A, Brown ; AL. et: Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. In: *GBD 2015 Disease Injury Incidence Prevalence Collaborators*, 2016

[2] T, Vos ; C, Allen ; M, Arora ; RM, Barber ; ZA, Bhutta ; A, Brown ; AL. et: 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. In: *GBD 2015 Mortality Causes of Death Collaborators*, 2016

[3] CP, Wild ; E, Weiderpass ; BW, Stewart: *World Cancer Report (PDF)*. World Health Organization, 2014. – pp. Chapter 5.14 S. – ISBN 978–9283204299

[4] *Melanoma Treatment -for health professionals.* `https://www.cancer.gov/types/skin/hp/melanoma-treatment-pdq`, 2015

[5] J, Daniel J. ; ELEWSKI, Boni E.: The ABCDEF Rule: Combining the ÄBCDE Ruleänd the Ügly Duckling Signïn an Effort to Improve Patient Self-Screening Examinations, The Journal of clinical and aesthetic dermatology, 2015

[6] *What is the abcde rule of skin cancer.* `https://www.outerbanksdermatology.com/blog/what-is-the-abcde-rule-of-skin-cancer`, 2023

[7] *Skin Cancer MNIST: HAM10000.* `https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000`,

[8] JAIN, Shivangi ; JAGTAP, Vandana ; PISE, Nitin: Computer Aided Melanoma Skin Cancer Detection Using Image Processing. In: *Procedia Computer Science* 48 (2015), S. 735–740. `http://dx.doi.org/https://doi.org/10.1016/j.procs.2015.04.209`. – DOI https://doi.org/10.1016/j.procs.2015.04.209. – ISSN 1877–0509

[9] PREMALADHA, J. ; RAVICHANDRAN, K.S: Novel Approaches for Diagnosing Melanoma Skin Lesions Through Supervised and Deep Learning Algorithms. In: *Journal of Medical Systems* 40 (2016). `http://dx.doi.org/10.1007/s10916-016-0460-2`. – DOI 10.1007/s10916–016–0460–2

[10] Thaajwer, MA. A. ; Ishanka, UA. P.: Melanoma Skin Cancer Detection Using Image Processing and Machine Learning Techniques. In: *2020 2nd International Conference on Advancements in Computing (ICAC)* Bd. 1, 2020, S. 363–368

[11] Ali, Abder-Rahman H. ; Li, Jingpeng ; Yang, Guang: Automating the ABCD Rule for Melanoma Detection: A Survey. In: *IEEE Access* 8 (2020), S. 83333–83346. http://dx.doi.org/10.1109/ACCESS.2020.2991034. – DOI 10.1109/ACCESS.2020.2991034

[12] Mabrouk, Mai S. ; Sayed, Ahmed Y. ; Afifi, Heba M. ; Sheha, Mariam A. ; Sharwy, Amr: Fully Automated Approach for Early Detection of Pigmented Skin Lesion Diagnosis Using ABCD. In: *Journal of Healthcare Informatics Research* 4 (2020), S. 151–173. http://dx.doi.org/https://doi.org/10.1007/s41666-020-00067-3. – DOI https://doi.org/10.1007/s41666–020–00067–3

[13] Márquez-Neila, Pablo ; Baumela, Luis ; Alvarez, Luis: A Morphological Approach to Curvature-Based Evolution of Curves and Surfaces. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36 (2014), Nr. 1, S. 2–17. http://dx.doi.org/10.1109/TPAMI.2013.106. – DOI 10.1109/TPAMI.2013.106

[14] Razmjooy, N. ; Mousavi, B. S. ; Soleymani, Fazlollah ; Khotbesara, M. H.: *A computer-aided diagnosis system for malignant melanomas*

[15] Jaworek-Korjakowska, Joanna: Computer-Aided Diagnosis of Micro-Malignant Melanoma Lesions Applying Support Vector Machines. In: *Hindawi Publishing Corporation* (2016). http://dx.doi.org/http://dx.doi.org/10.1155/2016/4381972. – DOI http://dx.doi.org/10.1155/2016/4381972