Interpretable Melanoma Detection for a Clinical Environment

Elliot Naylor

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

University of South Wales
Faculty of Mathematics and Computing

Date

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

Dedication here.

Acknowledgements

Abstract

Contents

List of Tables

Chapter 1

Background

Skin cancer is considered amongst the most severe public health concerns, with mortality rates of 2,353 per 100,000 within the United Kingdom (UK) in 2018 [32]. Skin cancers can be categorised between melanoma and non-melanoma, whereas melanoma is the most dangerous because it is unpredictable. When left untreated and after growing sufficiently, it can spread to other regions of the body (known as metastatic melanoma), which once progressed is challenging to treat effectively with a 10% survival over ten years in the US [7]. Furthermore, it is beneficial to catch melanoma early because it is the most easily treatable form of cancer, with 86% of cases being preventable [32]. However, melanoma can remain dormant from anywhere between 6 months to 10 years before maturing and becoming a danger to the patient [32]. Another danger of melanoma is its similarity to non-melanoma skin cancers, such as a mimic called seborrhoeic keratosis (SK), which frequently leads to misdiagnoses [14]. There are features unique to SK called fissures, ridges, and hairpin vessels [19]. Problematically these features require trained specialists to recognise them needing more than ten years of experience to have an accuracy of 86% compared to 62% or 56% (3-5 years of experience) [20]. However, because of the cost of training new doctors, there are limited available. Dermatologists primarily treat skin conditions (biopsies) and confirm diagnoses submitted by GP. General practitioners (GP) are the first to diagnose skin conditions and sometimes have limited experience diagnosing them (especially dermatological features). This project aims to improve the accuracy of GP observations by providing tools for the automatic classification of skin lesions. An automatic system should be cost-effective and advantageous to doctors.

Diagnostic procedures are instructions developed by doctors to simplify diagnosing conditions. Various methods have been developed to diagnose skin lesions and have greatly improved GP accuracy within clinical environments [15, 33]. Considering melanoma is the most dangerous skin condition, most procedures were developed specifically for early detection. Some include ABCD rules, 2-point checklist, 7-point checklist, and CASH. The most preferred of these techniques are ABCD rules and CASH because they have a higher

Criteria	Methodology	Score	Weight
Asymmetry	Measuring asymmetry involves first finding the centroid and splitting it twice with a 90-degree axis. Each side is subtracted with its opposite half to measure the asymmetry of shape, colour, and dermoscopic structures. If both sides are asymmetrical then the score is 2, one side asymmetrical is a score of 1, and otherwise, the score is 0.	0 - 2	× 1.3
Border	border is found by finding the centroid and drawing lines through it with a 45-degree angle, splitting the skin lesion into eight segments. Border segments might be irregular with convexity, sharp corners, or edges. Irregular segments are incremented by 1, reaching 8 for each segment.	0 - 8	× 1.3
Colour	The area of the skin lesion is up to 6 colours (white, red, light brown, dark brown, blue-grey, black). The score is increased by 1 for each visible colour, reaching a total of 6.	1 - 6	× 0.5
Dermoscopic Structures	Dermoscopic structures are measured by finding structureless areas, pigment networks, atypical networks, dots, and globules. Each visible structure adds a score of 1, reaching a total of 5.	1 - 5	× 0.5

Table 1.1: Total dermoscopy score (TDS) is a scoring system used with ABCD rules to support clinicians when diagnosing melanoma [8]. Each rule is multiplied by weights and the sum of the combined values is the final score, all together: [(Asymmetry × 1.3) + (Border × 0.1) + (Colour × 0.5) + (Dermoscopic structure × 0.5)].

sensitivity [33] and ABCD rules are generally the most preferred because it is easy to learn and is rapidly calculated [15]. There are several variations of ABCD rules, but, is originally measured using asymmetry, border, colour, and diameter. Diameter is sometimes replaced with dermatological structures because many features (i.e. blue-black signs, pigment networks, pseudopods, streaks, or milia-like cysts [28]) improve the classification accuracy between melanoma and the mimic seborrheic keratosis [8]. Furthermore, automatically measuring diameter is often difficult because it is dependent on the photo apparatus and the distance from the skin lesion, which is rarely consistent in research. Table 1 describes each rule in more detail, including a scoring system called total dermoscopy score (TDS), where each rule is assigned a score and combined to reach a result of either malignant, suspicious or benign. The criteria is: [(Asymmetry × 1.3) + (Border × 0.1) + (Colour × 0.5) + (Dermoscopic structure × 0.5)]. Each rule is calculated using the following descriptions in Table 1 and multiplied by their weight and then added together to reach a final score where [< 4.76 = benign, > 4.76 or < 5.45 = suspicious, > 5.45 = melanoma]. The disadvantage is the subjectivity of GP observations relating to their experience. So, it would be beneficial to automate the techniques using algorithms to standardise results and improve GP accuracy.

Computer-aided diagnostic (CAD) frameworks are a collection of algorithms designed to guide decision-making processes within clinical environments [9]. A paper written by Andre

Estava demonstrates a deep convolutional neural network (DCNN) that has comparable accuracy to that of dermatologists, trained using 129,450 clinical images consisting of 2,032 different diseases [5]. DCNN generates a collection of artificial neurons organised into layers, where each neuron receives input from a previous layer to perform a computation. The collection of layers is a network, which (once trained) ultimately measures the relationship between input parameters based on provided data. It is important to note that the accuracy is proportionate to the number of images and data quality for training that network. Unfortunately, these image samples are frequently private and unavailable to many institutions. Without adequate image data to test the capabilities of machine learning models, there is no method for measuring these biases and is therefore unsafe to use within clinical environments. Secondly, these approaches will often produce a parallel diagnosis, not explaining the results. There are many valuable techniques, but even the best techniques are inadequate for doctors without catering to interpretability. Other techniques are interpretable by considering diagnostic procedures, such as ABCD rules, many of which are described by Ali [3]. Techniques based on diagnostic procedures can be more easily tested for biases and provide further insight to GPs with the means to learn from and understand the results. Techniques include support vector machine (SVM), a supervised machine learning that uses regression analysis to categorise labelled data into two or more groups. The advantage means less data for training is required, and the model is interpretable.

Explainable AI (XAI) techniques have recently gained attention because the European general data protection regulation (GDPR and ISO/IEC 27001) stated that these approaches, commonly referred to as "black box" approaches, are challenging to utilise in medical environments. Since then, There has been significant progress in making neural network architectures more interpretable. A wide range of techniques [12, 26, 24] have since been developed, demonstrating that it is possible to make neural network techniques interpretable. However, the problem is instead the current scepticism on whether these techniques are trustworthy [31, 25], and they can produce realistic but incorrect results [13]. Some other interpretable techniques do not utilise neural networks. For example, Javier López-Labraca et al. [17] described an interpretable technique using multiple SVM models with colour and three dermoscopic structures (i.e., pigment networks, globules, and streaks). Bayesian fusion combines each model to calculate a diagnosis. Bayesian probability is a type of probability theory that uses probability distribution to estimate the values of unobserved variables. Bayesian fusion has a comparable accuracy to neural network techniques [29]. Therefore, overall results should be partially interpretable for use within clinical environments.

Doctors will often only have access to a patient for a short time before moving to another. CAD frameworks are beneficial because they speed up the process, can improve accuracy [9], and ensure the gathering of relevant data (ABCD rules). Furthermore, it could take days for a second opinion from another doctor, where an automatic system immediately provides it. Automated systems should also provide adequate explanations that can be understood quickly and easily by doctors. One method is to provide visual

explanations. Many authors [35, 16, 1] describe different ways to measure ABCD rules, including the asymmetry of skin lesions using bi-fold. Automated versions of the procedure use the centroid and moments of inertia to fold the skin lesion horizontally and vertically along the centroid. The overhung area on both axes is subtracted from the final score to measure asymmetrical or symmetrical. This technique produces an adequate visualisation that can provide GPs with an interpretable result. There is a range of other examples for ABCD rules, including border [16, 35, 2], colour [27, 30, 16], and dermoscopic structures [17] that use a range of interpretable algorithms that produce interpretable results.

Overall, many advanced machine learning techniques using neural networks lack the interpretability required within clinical environments. Furthermore, public datasets lack rarer skin conditions, making finding biases challenging. Automating the ABCD rules can solve this by using a technique that GPs are familiar with, and by using statistical models to extract relevant features (relating to the ABCD rules). This is followed by summarising rules using Bayesian fusion and calculating the significance of individual features.

1.1 Aim

- Develop an interpretable CAD framework based on the ABCD rules to diagnose skin lesions automatically. The goal is to utilise statistical models to extract each ABCD rule (asymmetry, border, colour, and Dermoscopic structure). Each rule will be trained using individual SVM models and are combined using Bayesian Fusion.

1.2 Objectives

- Develop and validate skin lesion segmentation and border cut-off approach for improved irregularity detection of ABCD rules using SegNet and LBPC.
- Develop and validate melanoma classification based on the diagnostic procedure ABCD rules (asymmetry, border, colour, and dermoscopic structures) for improved interpretability to doctors using various statistical techniques and SVM models.
- Develop and validate combining ABCD rules for the probabilistic analysis of the most dependent features using Bayesian fusion. This could include meta-data for gender, age, touch, feeling, and location on the body.

1.3 Contributions to knowledge

1. **Developing and validating a novel skin lesion segmentation approach for accurate border cut-off segmentation to improve border irregularity analysis using SegNet and LBPC.**

SegNet is highly accurate at finding the area for the segmentation of skin lesions but is inaccurate for measuring border irregularities because the border cut-off between skin and skin lesion is insufficient. Border irregularity detection necessitates an accurate cut-off for more reliable results, which SegNet does not provide. LBPC solves this problem by exaggerating the cut-off and improving the accuracy of border irregularity detection. However, the disadvantage of LBPC is its inaccuracy when finding the skin lesion area. By combining SegNet and LBPC, detecting the skin lesion area using SegNet, followed by adjusting the border with LBPC; retaining the accuracy of SegNet while improving the border cut-off accuracy. Experimental testing utilising the PH^2 dataset containing expert segmentation data will determine the benefits of segmentation.

2. **Developing and validating a novel asymmetry analysis approach for improved irregular asymmetry detection in skin lesions using moment-based texture analysis for improved bi-fold analysis and superpixels for improved asymmetry colour comparisons.**

   The disadvantage of asymmetry measuring techniques for skin lesions is rotational moments for creating bi-folds. Current bi-folds solely consider the silhouette of the skin lesion, with no consideration towards colour or texture. Furthermore, recent techniques have measured asymmetrical irregularities based on colour and texture. Producing a bi-fold based on the shape, colour, and texture using moment-based texture analysis should improve the accuracy of asymmetry detection. In addition, utilising superpixels to measure colour asymmetry to avoid merging important features improves accuracy. Both techniques will be validated using the PH^2 asymmetrical score.

3. **Developing and validating a novel interpretable melanoma classifier for improved interpretability of ABCD rules (asymmetry, border, colour, and dermoscopic structures) using feature extraction, support vector machines (SVM), and Bayesian fusion.**

   The disadvantage of many neural network-oriented techniques is their lack of adequate interpretability, making them challenging to utilise in clinical environments. However, ABCD rules (asymmetry, border, colour, and dermoscopic structures) are a diagnostic procedure that most doctors are familiar with; therefore, developing a system automating this procedure is beneficial. Feature extraction techniques aim to separate the data essential for each ABCD rule and train an SVM model from the extracted features. For example, bi-folds measure asymmetry, which can be modified to train an SVM model. Repeating this for border, colour, and dermoscopic structures ensures that each rule is independent. Finally, combining the Bayesian fusion results measures the probabilistic significance between ABCD rules and combines them into benign or malignant. Techniques will be validated using the PH^2 dataset for testing ABCD

rules and ISIC 2018 datasets for diagnosis.

4. **Developing and validating a novel interpretable melanoma classifier with meta-data including age, gender, feeling, and location on the body to improve classification accuracy between melanoma and seborrhoeic keratosis (SK) using Bayesian probability for a modifiable probabilistic analysis.**

   Seborrhoeic keratosis (SK) is a melanoma mimic because it sometimes shares clinical features with melanoma. Moreover, differentiating between the two with entirely image data can lead to inaccuracies. Including meta-data age, gender, feeling, and location on the body should improve accuracy because SK appears more frequently on the head or back of old male patients. Bayesian probability networks are considered highly modifiable and can generate results with incomplete input, meaning meta-data is only inputted when necessary, benefiting doctors and improving the diagnosis. The associated organisation has a vast amount of valuable meta-data alongside image data of skin lesions; a private dataset will be created from these results and used to validate results.

Chapter 2

Literature Review

2.1 Introduction

This chapter reviews statistical and machine learning algorithms for the automatic classification of ABCD rules and discussion on techniques beneficial for use within clinical environments.

2.2 Discussion

When doctors diagnose conditions using CAD, they should rationalise and build an explanation to prove their diagnosis, building criteria that other doctors understand. Currently, many techniques [5] called "black box" approaches produce parallel diagnosis that lacks an explanation. These are insufficient for use within some clinical environments for this reason. Instead, it would be beneficial for doctors to follow procedures they are familiar with, such as diagnostic procedures including ABCD rules. The reviewed techniques aim to automate the ABCD rules using various statistical and machine-learning techniques. Many are interpretable and suitable for clinical environments.

Hybrid machine learning techniques are recently gaining traction, an example by Ali combines results from both Gaussian naive Bayes (GNB) and a CNN [2] for border irregularity detection. The CNN ensures high-accuracy classification by finding the relationship between each component, and the GNB is interpretable. Results are combined using an ensemble approach, making a prediction probability. Such techniques are promising for use within clinical environments.

There is a lack of literature describing adequate visual representations for doctors, and it is understandable as there is still little evidence proving that CAD systems improve doctors decision making-processes [11]. It would be beneficial to create literature describing a catalogue of different visualisations that benefit doctors. Putting all this information together, alongside a questionnaire, might provide further insight into the visualisations that might be most useful to doctors.

2.2.1 Datasets

One fundamental problem is the overutilisation of private or privately annotated datasets, making a direct comparison of algorithms difficult; hence, this review has no direct comparisons. Some are between benign and malignant [18, 16, 2, 1] while others utilise private or never mention any datasets [16, 27, 30, 23, 35]. None compare their rules, likely because of subjectivity depending on the dermatologists that labelled them. More datasets should be public to assess individual rules and reach objective measurements. Until then, results conform with malignant, suspicious, or benign.

2.3 Previous Works on Automating ABCD Rules

Many CAD frameworks follow a methodology for the classification of skin lesions. These are listed below:

1. Segmentation – Image segmentation is the process of partitioning an image into multiple segments for more accessible analysis. These areas can be separated manually by a dermatologist (known as the ground truth) or separated automatically using statistical or machine learning algorithms.

2. Feature Extraction - Gathering features through filtering, morphology and other statistical approaches. ABCD rules include asymmetry, border, colour, and dermoscopic structures.

3. Combination - Combining the extracted features before using Principal Component Analysis (PCA) or after classification using Bayesian Fusion. Others combine the results using the Total Dermoscopy Score (TDS).

4. Classification – Measuring the results from the features and components through classification. Containing the final diagnosis of the type of skin lesion (Naveus, SK, or Melanoma)

2.3.1 Segmentation

Yading Yuan and Yeh Chi Lo describe a fully convolutional network (FCN) with an accuracy of 91.7% with the PH^2 dataset [34]. FCN is a variation of a CNN using 1x1 convolutions instead of dense layers. Essentially, an FCN forms a more complex function (generating a more complex neural network), whereas the CNN forms a less complex function, likely to degrade essential features. Therefore, more data is needed to train an FCN effectively than a CNN. After the convolution layers, transposed convolution layers (or deconvolution) and other layers (un-pooling) up-sample the input feature map to the size of the input image. Then, the network, trained from ground truth (human-generated segmentation mask) and the original images, can automatically generate segmentation masks based on textures and

colours of the skin lesion provided. There are dozens of examples of this, such as SegNet [6], which is another transposed CNN not designed initially for skin lesions but is effective at segmenting skin lesions.

E. Meskini et al. proposed using Otsu binarisation - a threshold technique that is effective at locating the border of a skin lesion after segmenting using Segnet [18]. Researchers proposed that when analysing the skin lesion border using ABCD rules, the original SegNet methods were ineffective because the ground truth is subjective - ineffective at finding the border cut-off between the skin lesion and skin. While SegNet has a 91.7% with the PH^2 dataset, the data is not effective at finding the precise border cut-off required for accurate border classification using ABCD rules. Therefore, researchers proposed the Otsu threshold to find the skin lesion border after segmenting using SegNet. Fan proposes another technique that uses a saliency-based segmentation approach to capture the area, followed by an Otsu threshold [10] to find the border cut-off from the skin lesion with a precision of 96.78% validated using the PH^2 dataset.

Pedro M.M. Pereira et al. proposed local binary pattern clustering (LBPC) to exaggerate the border, producing accurate results when classifying ABCD rules than ground-truth borders in the PH^2 dataset [22]. Local binary patterns (LBP) are texture descriptors calculated by comparing the centre pixel (of each pixel in the grey scaled image) with the eight neighbouring pixels as 'i', and converting it to a binary using the equation: $[if centroid > neighbour_i = 0, otherwise = 1]$. These eight neighbouring values produce a binary of 01101100 (decimal of 108) and change the centroid to 108. Next, the described process repeats on each other pixel in the image. Finally, the newly filtered image subtracted from the original grey-scaled image creates a segmentation mask with an accurate border cut-off. Finally, Pereira describes classification methods using SVM or FNN presenting the extracted border with an accuracy of 79% and 77% (respectively) with the MED-NODE dataset.

2.3.2 Handcrafted Features

Handcrafted features are the extraction of particular features using statistical algorithms—the benefit of separating data into components is a more accessible breakdown, improving explainability. In addition, this might instantiate trust for use within a clinical environment and prove more helpful to doctors.

Asymmetry

Asymmetry can be measured using the bi-fold technique, which involves drawing a line down the middle of the skin lesion and comparing the two halves to confirm whether the sides match, on both the horizontal and vertical axes, as shown in 2.3.2. If the two sides are greatly different, it could be a warning sign of melanoma. Asymmetry can be measured using the shape [35], colour[16] and texture [1].

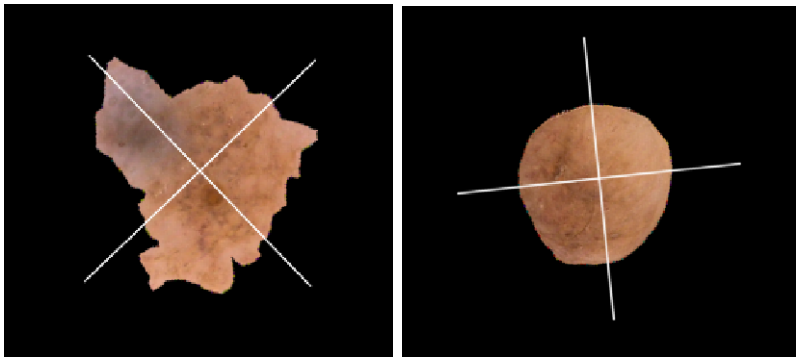

Figure 2.1: Images of two skin lesions from the PH^2 dataset showing the asymmetry calculated from moments.

Measuring the asymmetrical shape requires a precise border cut-off. Ihab S. Zaqout [35] describes a technique using the centroid and rotation of the skin lesion using moments of inertia. By Folding the skin lesion on both vertical and horizontal axes subtracting the opposite half. Pixels that cannot subtract are summed and compared with a threshold considering the skin lesion asymmetrical if the combined sum is more than the threshold.

Reda Kasmi and Karim Mokrani [16] describe creating a grid of 20x20 pixels of the skin lesion image and converting it into the LAB colour space. Next, each block's average colour is compared with a perpendicular block (vertical and horizontal axes) using the three-dimensional Euclidean luminance distance, a-axis, and b-axis. If more than half of the colour comparisons are over the threshold, that axis is considered colour asymmetrical. Blocks that have no symmetrical pair are ignored. Finally, luminance calculated separately prevents brightness problems. This technique has an accuracy of 94% with a private dataset.

Measuring similarities in texture can be achieved by using SIFT-based similarity and projection profiles [1]. SIFT is scale-invariant and helpful for texture components with varying texture quality. First, the skin lesion is split vertically and horizontally across the centre into four halves, comparing texture components on the symmetrical halves and measuring the similarity. Lastly, the projection profile in the x and y directions generates histograms. These results train a decision tree and have an 80% accuracy of the ISIC 2018 with 204 images privately annotated for ABCD rules and combined.

Border

Estimating border irregularities involves splitting the skin lesion into eight equal sections (through the centroid), where each section with tight corners and convexity is considered irregular. Each irregular section of the border adds a score of 1 ranging from 0 to a total of 8, as shown in figure 2.3.2.

Border irregularity contours were found by splitting the skin lesion into eight segments

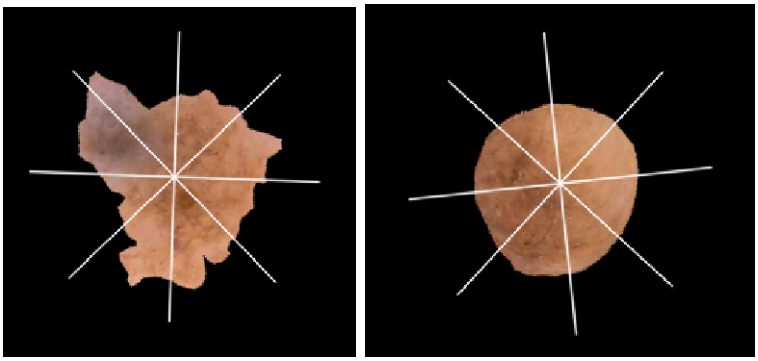

Figure 2.2: Images of two skin lesions split into 8 sections using moments, each border is measured for irregularity.

around the centre, and then calculating a fitting error for each. If the error is larger than 0.05 (x contour), that area is considered irregular [16].

Abder Rahman H. Ali et al. calculate the compactness of each border by first calculating the contour around the area of the lesion containing x and y positions. Next, measure the space between each position to estimate the compactness. The tighter the curves and corners, the more contour positions, revealing irregular borders within a segment, combining all of these scores creates the irregularity index [35].

Fractal dimensions (FDs) is a statistical index measuring the detail in a pattern changing with the image scale index. One technique called box-counting increases values if there are more corners and edges around the border. The higher the value demonstrates the level of border irregularity. Ali describes using machine learning alongside Zernike moments, and convexity measurements for a high-accuracy border irregularity classification [2]. However, results are ambiguous because the output is either "irregular" or "regular" border (not relating to the TDS). Thus, conforming to the TDS and splitting the border into eight sections would make it more interpretable and useful o doctors. However, a hybrid GNB and CNN approach are combined to allow interpretability through GNB.

Colour

Colour refers to the shades of pigment within the area of a skin lesion, not referring to abnormalities relating to bruises, crust, and grazes. Melanoma usually contains more than two colours compared with benign lesions, singular in colour. Skin lesions can consist of one or many colours: white, red, light brown, dark brown, blue-grey and black.

Finding colour variations has been achieved by calculating the normalised standard deviation of the red, green, and blue components [27]. The normalisation process improves the recognition of normal skin pigmentation, which would show pigmentation levels, making comparisons easier between different skin lesions.

Arthur Tenenhaus, et al. utilise joint learning using Kohonen map, and k-means clustering [30]. Five random pixels create a 5x5 Kohonen map represented by 25 neurons in a neural network for each skin lesion in the dataset. Colour variations on a 25-dimensional vector find the proportions of pixels projected onto each of the 25 neurons. Next, K-means classifies the skin lesions set by the number of colours found by dermatologists. Only four colours were present in the dataset in this scenario, while seven could be. Eventually, the colour components are represented as a 42-dimensional vector and are passed into a KL-PLS based classifier to detect variations in colour at 66% using a private dataset.

Reda Kasmi et al. locate the number of colour variations by converting the image into the LAB colour space matching the colour ranges that can be perceived by human eyes [21], measuring the average colour distribution of the dataset and assigning each colour as a threshold range. Next, the Euclidean distance between each colour threshold is compared with each pixel colour [16], finding the closest matching colour of the six colours. Finally, removing the areas of colour with less than 5% prevents the classification of dots. This approach uses a colour range of white, light brown and dark brown. However, there is a static threshold value for the other colours, which would be unlikely to cover the ranges of the colours, including red, blue-grey, and black.

Dermoscopic structures

Dermoscopic structure refers to structures on the skin lesion, including pigment networks, structureless areas, dots, globules, streaks, white structures, and 22 others (not including sub-types). Variations of pigment networks are more commonly found in melanoma [4] and are therefore a valuable feature for automatic classification. Similarity other features such as milia-like cysts, a sub-type called milia-like cysts (MLC) called cloudy MLC appears more frequently on melanoma than SK, with a specificity of 99.1% specificity [28].

Javier López-Labraca et al. [17] describes a statistical approach to classifying melanoma using dermoscopic structures through Gabor filtering, support vector machines, and Bayesian fusion. This technique uses a form of soft segmentation to find the area of these dermoscopic features. Firstly the structures are located using Gabor filtering using different values to find fissures and globules. Each structure is then compared with a trained SVM model to check the similarity of the detected features. The results from the model are then combined using Bayesian fusion to reach a result of malignant or benign. Finally, training a CNN model alongside an SVM improves the retractability of dermoscopic structures; compared to a standalone CNN model.

Combining ABCD Rules

This section describes combining features from the ABCD rules into a classification between malignant, suspicious or benign after considering all clinical features. Again, meta-data and texture can potentially improve the results.

Maryam Ramezani et al. proposed a method to extract features from ABCD rules storing them in vectors and extracting the texture as a GLCM. First, these 187 features are shrunk to 13 using PCA [23]. Next, the data trains an SVM to classify skin lesions into benign or malignant with an accuracy of 82.2% on macroscopic images using a private dataset.

Other methods output TDS [35, 36], which combines them using: [(Asymmetry x 1.3) + (Border x 0.1) + (Colour x 0.5) + (Diameter x 0.5)]. A statistical model for each ABCD rule outputs a score in the same format. The benefit is interpretability because it follows the diagnostic procedure. The technique achieved an accuracy of 90% using a private dataset.

2.4 Conclusion

Many techniques utilise ABCD rules to produce an automatic and interpretable diagnosis. Interestingly, many focus on detecting and classifying asymmetry, border, and colour (ABC) or dermoscopic structures, but neither combine the whole ABCD rules into a single framework. Despite dermoscopic structures providing a means of diagnosing problematic forms of melanoma, including mimics (seborrhoeic keratosis) [14], and non-pigmented melanomas. Thus, it would be valuable to combine both into a single system for possibly higher accuracy.

Despite various valuable features, asymmetry rarely utilises techniques other than statistical models. For example, researchers highly focused on border irregularity and dermoscopic structures, leading to hybrid machine-learning models for their assessment. However, asymmetry still utilises statistical approaches to measure and combine shape, colour, and texture. It would be beneficial to transform this data and process it using an SVM, improving accuracy.

Utilising external data, including feeling, touch, age, and location on the body, are helpful to doctors when diagnosing skin conditions, but is not mentioned in any of the discussed techniques. It would be beneficial to implement this data into the decision-making process.

Chapter 3

Data Extraction and Transformation

Chapter 4

Proposed Method

Bibliography

[1] Abder Rahman Ali, Jingpeng Li, and Sally Jane O'Shea. "Towards the automatic detection of skin lesion shape asymmetry, color variegation and diameter in dermoscopic images". In: *PLoS ONE* 15.6 (2020). ISSN: 19326203. DOI: 10.1371/journal.pone.0234352.

[2] Abder-Rahman Ali et al. "A machine learning approach to automatic detection of irregularity in skin lesion border using dermoscopic images". In: *PeerJ Computer Science* 6 (2020), e268. ISSN: 2376-5992. DOI: 10.7717/peerj.cs.268/table-3.

[3] Abder Rahman H. Ali, Jingpeng Li, and Guang Yang. "Automating the ABCD Rule for Melanoma Detection: A Survey". In: *IEEE Access* 8 (2020), pp. 83333–83346. ISSN: 21693536. DOI: 10.1109/ACCESS.2020.2991034.

[4] Murali Anantha, Randy H. Moss, and William V. Stoecker. "Detection of pigment network in dermatoscopy images using texture analysis". In: *Computerized Medical Imaging and Graphics* 28.5 (2004). ISSN: 08956111. DOI: 10.1016/j.compmedimag.2004.04.002.

[5] Esteva Andre et al. "Dermatologist-level classification of skin cancer with deep neural networks". In: *Nature* 542.7639 (2017), pp. 115–118. ISSN: 1476-4687. DOI: 10.1038/nature21056LK-http://elinks.library.upenn.edu/sfx_local?sid=EMBASE&issn=14764687&id=doi:10.1038%2Fnature21056&atitle=Dermatologist-level+classification+of+skin+cancer+with+deep+neural+networks&stitle=Nature&title=Nature&volume=542&issue=7639&spage=115&epage=118&aulast=Esteva&aufirst=Andre&auinit=A.&aufull=Esteva+A.&coden=NATUA&isbn=&pages=115-118&date=2017&auinit1=A&auinitm=. URL: http://www.embase.com/search/results?subaction=viewrecord{\&}from=export{\&}id=L614981551{\%}0Ahttp://dx.doi.org/10.1038/nature21056.

[6] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.12 (2017), pp. 2481–2495. ISSN: 01628828. DOI: 10.1109/TPAMI.2016.2644615. arXiv: 1511.00561.

[7] Shailender Bhatia, Scott S. Tykodi, and John A. Thompson. *Treatment of metastatic melanoma: An overview*. 2009.

[8] Armand B. Cognetta et al. "The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions". In: *Journal of the American Academy of Dermatology* 30.4 (1994), pp. 551–559. ISSN: 01909622. DOI: 10.1016/S0190-9622(94)70061-3.

[9] Vincent Dick et al. "Accuracy of Computer-Aided Diagnosis of Melanoma: A Meta-analysis". In: *JAMA Dermatology* 155.11 (2019), pp. 1291–1299. ISSN: 21686068. DOI: 10.1001/jamadermatol.2019.1375.

[10] Haidi Fan et al. "Automatic segmentation of dermoscopy images using saliency combined with Otsu threshold". In: *Computers in Biology and Medicine* 85 (2017), pp. 75–85. ISSN: 18790534. DOI: 10.1016/j.compbiomed.2017.03.025.

[11] Lavinia Ferrante di Ruffano et al. "Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults". In: *Cochrane Database of Systematic Reviews* 12 (2018). ISSN: 1469493X. DOI: 10.1002/14651858.CD013186.

[12] Masaru Fuji et al. "Explainable AI through combination of deep tensor and knowledge graph". In: *Fujitsu Scientific and Technical Journal* 55.2 (2019), pp. 58–64. ISSN: 00162523.

[13] Amirata Ghorbani, Abubakar Abid, and James Zou. "Interpretation of Neural Networks Is Fragile". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (2019), pp. 3681–3688. ISSN: 2159-5399. DOI: 10.1609/aaai.v33i01.33013681. arXiv: 1710.10547.

[14] Leonid Izikson et al. "Prevalence of melanoma clinically resembling seborrheic keratosis: Analysis of 9204 cases". In: *Archives of Dermatology* 138.12 (2002). ISSN: 0003987X. DOI: 10.1001/archderm.138.12.1562.

[15] Reda Kasmi and Karim Mokrani. "Classification of malignant melanoma and benign skin lesions: Implementation of automatic ABCD rule". In: *IET Image Processing* 10.6 (2016), pp. 448–455. ISSN: 17519659. DOI: 10.1049/iet-ipr.2015.0385.

[16] Reda Kasmi and Karim Mokrani. "Classification of malignant melanoma and benign skin lesions: Implementation of automatic ABCD rule". In: *IET Image Processing* 10.6 (2016), pp. 448–455. ISSN: 17519659. DOI: 10.1049/iet-ipr.2015.0385.

[17] Javier López-Labraca et al. "Enriched dermoscopic-structure-based cad system for melanoma diagnosis". In: *Multimedia Tools and Applications* 77.10 (2018), pp. 12171–12202. ISSN: 15737721. DOI: 10.1007/s11042-017-4879-3.

[18] E. Meskini et al. "A new algorithm for skin lesion border detection in dermoscopy images". In: *Journal of Biomedical Physics and Engineering* 8.1 (2018), pp. 109–118. ISSN: 22517200. DOI: 10.22086/jbpe.v0i0.444.

[19] Akane Minagawa. "Dermoscopy–pathology relationship in seborrheic keratosis". In: *Journal of Dermatology* 44.5 (2017), pp. 518–524. ISSN: 13468138. DOI: 10.1111/1346-8138.13657.

[20] C. A. Morton and R. M. Mackie. "Clinical accuracy of the diagnosis of cutaneous malignant melanoma". In: *British Journal of Dermatology* 138.2 (1998). ISSN: 00070963. DOI: 10.1046/j.1365-2133.1998.02075.x.

[21] Nikolaos E. Myridis. "Ultra-realistic Imaging: Advanced Techniques in Analogue and Digital Colour Holography, by Hans Bjelkhagen and David Brotherton-Ratcliffe". In: *Contemporary Physics* 55.3 (2014), pp. 247–248. ISSN: 0010-7514. DOI: 10.1080/00107514.2014.907348.

[22] Pedro M.M. Pereira et al. "Skin lesion classification enhancement using border-line features – The melanoma vs nevus problem". In: *Biomedical Signal Processing and Control* 57 (2020). ISSN: 17468108. DOI: 10.1016/j.bspc.2019.101765.

[23] Maryam Ramezani, Alireza Karimian, and Payman Moallem. "Automatic Detection of Malignant Melanoma using Macroscopic Images". In: *Journal of Medical Signals and Sensors* 4.4 (2014), pp. 281–290. ISSN: 22287477. DOI: 10.4103/2228-7477.144052.

[24] Marco Ribeiro, Sameer Singh, and Carlos Guestrin. ""Why Should I Trust You?": Explaining the Predictions of Any Classifier". In: *ArXiv* (2016), pp. 97–101. DOI: 10.18653/v1/n16-3020.

[25] Wojciech Samek and Klaus Robert Müller. "Towards Explainable Artificial Intelligence". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11700 LNCS (2019), pp. 5–22. ISSN: 16113349. DOI: 10.1007/978-3-030-28954-6_1. arXiv: 1909.12072.

[26] Ramprasaath R. Selvaraju et al. "Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization". In: *Revista do Hospital das Cl??nicas* 17 (2016), pp. 331–336. ISSN: 00418781. arXiv: 1610.02391. URL: http://arxiv.org/abs/1610.02391.

[27] Zhishun She, Y. Liu, and A. Damatoa. "Combination of features from skin pattern and ABCD analysis for lesion classification". In: *Skin Research and Technology* 13.1 (2007), pp. 25–33. ISSN: 0909752X. DOI: 10.1111/j.1600-0846.2007.00181.x.

[28] S. M. Stricklin et al. "Cloudy and starry milia-like cysts: How well do they distinguish seborrheic keratoses from malignant melanomas?" In: *Journal of the European Academy of Dermatology and Venereology* 25.10 (2011), pp. 1222–1224. ISSN: 09269959. DOI: 10.1111/j.1468-3083.2010.03920.x.

[29] Maen Takruri and Abubakar Abubakar. "Bayesian decision fusion for enhancing melanoma recognition accuracy". In: *2017 International Conference on Electrical and Computing Technologies and Applications, ICECTA 2017* 2018-Janua (2017), pp. 1–4. DOI: 10.1109/ICECTA.2017.8252063.

[30] Arthur Tenenhaus et al. "Detection of melanoma from dermoscopic images of naevi acquired under uncontrolled conditions". In: *Skin Research and Technology* 16.1 (2010), pp. 85–97. ISSN: 0909752X. DOI: 10.1111/j.1600-0846.2009.00385.x.

[31] Erico Tjoa and Cuntai Guan Fellow. "A Survey on Explainable Artificial Intelligence (XAI): Towards Medical XAI". In: *arXiv* (2019). ISSN: 23318422. DOI: 10.1109/tnnls.2020.3027314. arXiv: 1907.07374. URL: http://arxiv.org/abs/1907.07374.

[32] Cancer Research UK. *Melanoma skin cancer statistics — Cancer Research UK*. 2019. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/melanoma-skin-cancer{\#}heading-Three.

[33] Ezgi Unlu, Bengu N. Akay, and Cengizhan Erdem. "Comparison of dermatoscopic diagnostic algorithms based on calculation: The ABCD rule of dermatoscopy, the seven-point checklist, the three-point checklist and the CASH algorithm in dermatoscopic evaluation of melanocytic lesions". In: *Journal of Dermatology* 41.7 (2014), pp. 598–603. ISSN: 13468138. DOI: 10.1111/1346-8138.12491.

[34] Yading Yuan and Yeh Chi Lo. "Improving Dermoscopic Image Segmentation With Enhanced Convolutional-Deconvolutional Networks". In: *IEEE Journal of Biomedical and Health Informatics* 23.2 (2019), pp. 519–526. ISSN: 21682194. DOI: 10.1109/JBHI.2017.2787487. arXiv: 1709.09780. URL: http://arxiv.org/abs/1703.05165{\%}0Ahttp://dx.doi.org/10.1109/JBHI.2017.2787487.

[35] Ihab S. Zaqout. "Diagnosis of Skin Lesions Based on Dermoscopic Images Using Image Processing Techniques". In: *International Journal of Signal Processing, Image Processing and Pattern Recognition* 9.9 (2016), pp. 189–204. ISSN: 20054254. DOI: 10.14257/ijsip.2016.9.9.18.

[36] Yongfeng Zhang and Xu Chen. "Explainable Recommendation: A Survey and New Perspectives". In: (2018). arXiv: 1804.11192. URL: http://arxiv.org/abs/1804.11192.

()

Chapter 5

Tables