SAA: A novel skin lesion Shape Asymmetry Classification Analysis

Shaik Reshma¹ and Reeja S R²,*

¹,²School of Computer Science and Engineering, VIT-AP University

Abstract

INTRODUCTION: Skin cancer is emerging as a significant health risk. Melanoma, a perilous kind of skin cancer, prominently manifests asymmetry in its morphological characteristics.

OBJECTIVE: The objective of the study is to classify the asymmetry of the skin lesion shape accurately and to find the number of symmetric lines and the angles of formation of symmetric lines.

METHOD: This study introduces a unique methodology known as Shape Asymmetry Analysis (SAA). The SAA incorporates a comprehensive framework including image pre-processing, segmentation along with the computation of mean deviation error and the subsequent categorization of data into symmetric and asymmetric forms using a classification model.

RESULT: The PH2 dataset is used in this study, where the three labels are consolidated into two categories. Specifically, the labels “symmetric” and “symmetric with one axis” are merged and classified as "symmetric," while the label “asymmetric” is unchanged and classified as "asymmetric". The model demonstrates superior performance compared to conventional methodologies, achieving a noteworthy accuracy rate of 90%. Additionally, it exhibits a weighted F1-score, precision, and recall of 0.89, 0.91, 0.90 respectively.

CONCLUSION: The SAA model accurately classifies skin lesion shapes compared to state-of-the-art methods. The model can be applied to the shapes, irrespective of irregularity, to find symmetric lines and angles.

Keywords: Shape Asymmetry, Mean Deviation Error, Skin Surface Microscopic Images

Received on 25 December 2023, accepted on 22 March 2024, published on 28 March 2024

Copyright © 2024 S. Reshma et al., licensed to EAI. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetpht.10.5580

1. Introduction

Cancer is a group of disorders characterized by unregulated and abnormal cell growth inside the human body. In a condition of peak physical health, cellular processes such as growth, division, and programmed cell death are carefully managed, therefore playing a role in maintaining the normal physiological functions of tissues and organs. However, in the context of cancer, this biological process becomes aberrant. Cell division and proliferation occur in an excessive manner, resulting in the formation of an aggregate of tissue frequently denoted as a neoplasm. Skin cancer is a malignancy that arises from the dermal cells of the integumentary system. The most prevalent form of cancer is attributed primarily to the effects of ultraviolet (UV) radiation, originating from either natural sources such as the sun or artificial ones like tanning beds [1].

Melanoma and non-malignant melanoma, which include basal and squamous cell carcinomas, are the two principal forms of skin cancer [2]. The diagnosis of skin malignancy encompasses both invasive and non-invasive techniques. The invasive procedure of biopsy is employed to identify the presence of malignancy, albeit at the cost of inducing pain and discomfort for the patient as well as necessitating a significant time commitment. Digital skin-surface microscopy is a commonly accepted non-surgical technique that is considered to be cost-effective, minimally painful, and efficient for identifying and categorizing skin cancer. The
ABCD rule checklist [3], which was established by a troop of scientists, is a well-regarded non-invasive method used for the timely identification of possibly malignant pigmented lesions. Among all the rules, the form of the lesion is an important aspect to examine, namely A-form asymmetry, B-Irrregularity of borders, C-Lesion Colour change, and D-Lesion Diameter. The existence of asymmetry plays an important contribution in the assessment of the malignancy of a lesion. Non-malignant tumors typically exhibit circular and symmetrical shapes. In contrast, melanoma presents as an unregulated proliferation and disseminates in an asymmetrical manner [4]. In the domain of dermatology, an asymmetric configuration pertains to skin lesions or moles that lack symmetry, wherein one side of the lesion does not exhibit a mirrored or congruent counterpart when wrapped along its main axis, meaning the two portions of the lesion exhibit dissimilarities in both shape and appearance [5].

The Ph2 dataset [6], which is publicly accessible, is used for experimental purposes due to its inclusion of accurate values regarding the asymmetry of skin surface malignant shape images. The 200 images in dataset have been classified into three distinct categories based on their degree of asymmetry: 0, which signifies complete symmetry; 1, indicating symmetry with respect to a single axis; and 2, representing full asymmetry. Figure 1 depicts images of skin lesions belonging to various categories.

![Figure 1. Instance images from PH2 dataset (a) Symmetric with one axis, (b) Fully Asymmetric, (c) Fully symmetric.](Image)

This paper focuses on:

- The SAA model aims to identify skin lesion asymmetry.
- The extraction of a skin lesion segmentation mask for the purpose of identifying asymmetry. The widely used medical image segmentation model, U-Net, is employed for this task.
- The mid values of skin lesion shape are extracted and it is observed that they are non-linear in nature, which is attributed to the irregular contours of the lesions.
- To measure the asymmetry of the skin surface malignant shape, a straight line is drawn passing through the majority of the mid values, and the deviation error is determined.
- Skin lesions are classified as symmetric or asymmetric using the most effective machine learning model.

2. Literature Survey

Talavera-Martinez et al. [7] used deep learning to diagnose skin lesions by symmetry. This is a CNN-based model has 10 layers. An input layer, 3 convolution and down-sampling blocks, 2 dense layers, and a Softmax layer follow. The Softmax layer outputs a 3-class categorization based on lesion symmetry. The experiment used 1053 pictures. The PH2 dataset yielded 438 pictures with ground truth values. The remaining 615 photos were from 262 EDRA 2002, 177 ISIC, 41 Dermis, and 135 Dermquest skin cancer databases.

Three domain experts used majority vote to annotate 615 lesions. The strategy provided in the study demonstrated superior performance compared to conventional methods and trained beforehand networks. It achieved an F1-score, accuracy, sensitivity, and specificity of 64.5%, 70.7%, 63.6%, and 73.5%, respectively.

Toureau et al. [8] developed computational algorithms to determine symmetry in dermoscopic skin lesion images. This work presents three computational approaches for measuring skin lesion symmetry: shape, colour, texture, and their combined features. Original method focuses exclusively on lesion form and parameterizes possible symmetry axes, notably lines across the center of mass. Jaccard index assesses lesion axis symmetry. Aggregating shape-based symmetry indices using a random forest classifier classifies input as “no symmetry,” “symmetry with one axis,” or “symmetry with two axes.” The second method compares dermoscopic image textures using a unique dataset to highlight the lesion's chromatic and textural qualities. The third method improves symmetry identification by combining the first two. The skin lesion symmetry detection technique is tested on ph2. The study indicated that the method identifies skin lesion symmetry with 84%–88% accuracy.

The shape asymmetry quantifier Fourier descriptors were used by Clawson et al. [9] to identify asymmetry in skin lesions. Fourier descriptors may be used in diagnostic processes to precisely identify and prioritize lesion symmetry axes. The study uses 30 dermoscopic pictures of pigmented skin lesions with varied asymmetry. Manually collected lesion boundary points totally 64 per lesion. The major axis of symmetry for each lesion was determined by two dermatologists who evaluated them blindly. The minimal % asymmetry of each lesion was estimated using a folding method. This included determining the mean difference between A min and the primary Fourier axis percentage asymmetry. The Fourier approach for identifying symmetry axes is exact, robust, and correlated with clinical diagnosis. This research is limited by its tiny test set of 30 photos.

Ruela et al. [10] examined how form and symmetry affect dermoscopic melanomas detection. The researchers created a CAD system that splits the picture in half along the primary axes. Subsequently, one portion of the lesion is folded in a manner that duplicates the other, resulting in the establishment of fundamental symmetrical attributes. The system then calculates symmetry. The CAD system used form characteristics with 83% sensitivity and 78% specificity. A 96% sensitivity and 86% specificity on symmetry characteristics were used on the PH2 dataset. The authors also compare form and symmetry and colour and texture outputs using the same CAD technology.

in dermoscopic pictures. These traits indicate malignant melanoma. Experimental photos came from the ISIC repository. A decision tree system detects skin lesion asymmetry with 80% accuracy. Extraction of colour variegation features detected suspicious colours. The ferret diameter measures the skin lesion's diameter. Finally, the Otsu-II algorithm segments images with an 87.7% dice coefficient.

Sirakov et al. [11] developed a technique and software application to automate skin lesion border, symmetry, and area extraction. LBESA uses image enhancement, S-ACES, and symmetry assessment. The suggested approach uses image enhancement, active contour evolution, and minimum boundary box computations to properly depict lesions as discrete points in a two-dimensional area-symmetry space. Using statistical analysis, 51 skin lesion photos were used to confirm theoretical ideas and evaluate border extraction accuracy against the ground truth. The LBESA system automatically identifies areas 95% accurately. The tool found that this methodology's S-ACES algorithm is fast and accurate. However, this method may struggle with high noise levels. The technology also suggests that LBESA may not work well for severe cancer pictures. Table 1 represents the pre-existing models summary.

Table 1. Results from pre-existing models.

<table>
<thead>
<tr>
<th>Author &amp; Reference</th>
<th>Methodology</th>
<th>Dataset</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talaver et al [7]</td>
<td>CNN Deep learning model</td>
<td>Own dataset with 615 images</td>
<td>Accuracy = 70.7% F1-score = 64.5% Sensitivity = 63.6% Specificity = 73.5%</td>
</tr>
<tr>
<td>Toureau et al [8]</td>
<td>Symmetry detection based on shape, colour and texture</td>
<td>PH2 dataset</td>
<td>Accuracy = 84-88%</td>
</tr>
<tr>
<td>Ruela et al [10]</td>
<td>Image split in half and overlapped</td>
<td>PH2 dataset</td>
<td>Sensitivity = 83% Specificity = 78%</td>
</tr>
<tr>
<td>Sirakov et al [11]</td>
<td>LBESA</td>
<td>51 dermal images</td>
<td>Accuracy = 96% Precision = 0.99 Recall = 0.75 Border error = 5.4%</td>
</tr>
</tbody>
</table>

3. Methodology

The proposed methodology, referred to as Shape Asymmetry Analysis (SAA), has been created to help medical professionals analyse skin lesion asymmetry. The classification of the asymmetry is accomplished by the use of technologies such as Machine Learning, Deep Learning and Image Processing. The SAA's detailed architecture is shown in Figure 2. Algorithm 1 represents the steps in SAA model.

3.1. Image pre-processing

The first and foremost phase is the pre-processing of the raw photos obtained from the dataset. The primary objective is to reduce the dimensions of the image to 256x256 in order to facilitate its use as input in deep learning models. Furthermore, the de-noising process involves the elimination of superfluous pixels, such as hair seen in photos of skin. In this study, the elimination of artefacts is achieved by a series of filters, starting with a morphological median filter and concluding with a bottom hat filter and harmonic in-painting technique [12].

3.2. Image segmentation

In this study, a U-Net model [13] that has been optimized is used to extract a binary mask from pictures that portray skin lesions. This architectural approach incorporates pixel-wise classification via the integration of contracting and expanding routes. The contracting pathway comprises layers such as the convolution operation, rectified linear activation function (ReLU), max-pooling operation, and down sampling. The process of expanding to up sampling involves the use of concatenation computation, convolution, and rectified linear activation function procedures. In the last layer, a solitary convolutional layer is used. The pre-processed pictures are then sent into the U-Net architecture to provide segmentation results. These segmented images are then used to calculate the degree of asymmetry.

3.3. Proposed framework

The white area of the segmented binary picture represents the lesion. Figure 3 represents the architecture of the proposed framework in the SAA model.

Locating centre

The first stage is the determination of the centroid of the white area inside the ROI by using the bounding boxes. A rectangular bounding box inside which the lesion area will be precisely accommodated. The coordinates x, y, height, and width are derived from the computation of the bounding box. The equation 1,2 is used to get the coordinates of the central point (x_c, y_c).

\[ x_c = x + \text{width}/2 \]
\[ y_c = y + \frac{\text{height}}{2} \]  

**Drawing base horizontal and parallel lines**  
Drawing a horizontal line across the spot, using a base \( x \)-coordinate of 0, as a base horizontal line using equation 3. A few lines that are parallel to the horizontal base line are drawn, maintaining a consistent distance between each line throughout the whole height of the lesion. The top portion of the lesion is delineated using equation 4, while the lower portion of the lesion is demarcated using equation 5, with reference to the base horizontal line.

\[ y = y_c \]  
\[ y_t = y_c - i \times \text{gap} \]  
\[ y_b = y_c + i \times \text{gap} \]  

**Locating parallel horizontal line midpoints**  
The midpoints of all the horizontal lines drawn on the lesion form are determined by evaluating the coordinates of the edge points \((x_1, y_1)\) and \((x_2, y_2)\) using the equation 6. Upon examination of the midpoints of all the horizontal lines, it becomes evident that they exhibit non-linear characteristics due to the irregularity in their form.

\[ \text{Midpoint} = \left( \frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right) \]  

**Drawing line across most midpoints**  
The RANSAC method is often used to estimate a straight line that best fits the majority of midpoints while excluding any non-linear points as outliers. This method operates on an iterative basis, with the number of iterations being determined by equations 7 and 8.

\[ N = \frac{\log(1-a)}{\log(1-(1-b^c))} \]  
\[ b = 1 - d \]  

**Calculating mean deviation error**  
The discrepancy between the linear equation represented as equation 9 and the midpoint values is quantified as the mean deviation error (MDE). Equation 10 is used for the computation of the MDE.

\[ y = A + Bx \]  
\[ MDE = \frac{\sum_{i=1}^{N} (y_i - A - Bx_i)^2}{N} \]  

**Rotating the lesion shape throughout 180 degrees**  
To ascertain the quantity of symmetric axes, the configuration of the lesion is subjected to a 180-degree rotation, and the angle at which zero mean deviation error appears is examined to detect the existence of symmetrical lines.

**3.4. Classification**  
The PH2 dataset consists of three distinct categories of ground truth labels: 0 represents symmetric lesions, 1 represents lesions that are symmetric with one axis, and 2
Figure 3. The architecture of proposed framework in SAA model.

represents lesions that are asymmetric. In this experiment, the ground truth values are transformed by consolidating three labels into two categories: symmetric and asymmetric. The labels "symmetric" and "symmetric with 1 axis" are merged into the category "symmetric," but the label "asymmetric" remains unchanged. Support Vector Machine (SVM) [15] is used to categories symmetric and asymmetric skin lesions. The MDE was computed for all 180-degree angles, along with the real labels passed as parameters to the SVM model.

Algorithm 1: Shape Asymmetry Analysis
1. Resize the input images.
2. Remove artefact hair
3. Create binary segmented image
4. Find centroid \( \left( x_c, y_c \right) \) of the skin lesion area.
5. While \( \text{angle} < 180 \) degrees
   - Draw horizontal line throughout the skin lesion area
   - While line! = Empty
     - Extract line segment end points \((x_2, y_2)\)
     - Mid-point \( \left( \frac{x_1+x_2}{2}, \frac{y_1+y_2}{2} \right) \)
   - End while
   - Input all mid points to RANSAC
   - Calculate MDE
   - Rotate lesion shape with a degree
6. End while
7. Apply classifier

4. Experimental Analysis

4.1. Dataset

The Ph2 dataset comprises 200 pigmented lesion images, each with an average dimension of 768x560 pixels in colour. These images are categorized into three groups based on skin cancer types: nevi, atypical nevi, and melanoma. Additionally, this dataset includes clinical annotations for all the lesion pictures, encompassing various aspects such as binary mask, morphological irregularity, reticular pattern, nodules, veins, regression regions, blue whitish veil, and lesion heterochromia. The ground truth for lesion asymmetry is provided in three values: 0 for complete symmetry, 1 for symmetry with respect to a single axis, and 2 for full asymmetry.

4.2. Performance outcome

The effectiveness of the SAA is evaluated via the use of many performance indicators, including accuracy, F1-score, precision, recall, and the ROC curve. The assessment of segmented pictures is quantified by using accuracy and IOU metrics. Equations 11, 12, 13, 14, 15 represents the formula to calculate performance indicators.

\[
\text{Accuracy} = \frac{\text{TruePositive} + \text{TrueNegative}}{\text{TruePositive} + \text{TrueNegative} + \text{FalsePositive} + \text{FalseNegative}} \tag{11}
\]

\[
F_1 - \text{SCORE} = \frac{2\text{TruePositive}}{2\text{TruePositive} + \text{FalsePositive} + \text{FalseNegative}} \tag{12}
\]

\[
\text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}} \tag{13}
\]

\[
\text{Recall} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}} \tag{14}
\]
\[ \text{IoU} = \frac{\text{Area of overlap}}{\text{Area of union}} \]  

(15)

4.3. Segmentation output

The U-Net model has been trained for a total of 30 epochs, resulting in an accuracy of 95.1% and IOU value of 0.16. Figure 4 displays the segmentation mask attained by the U-Net model in comparison to the ground truth values, demonstrating its efficiency.

![Figure 4: Segmentation result.](image)

4.4. Geometric shape outputs

In order to validate the SAA model for symmetric forms, numerous geometrical shapes with inherent symmetry were examined in conjunction with their corresponding symmetric axes. According to Figure 5, it can be seen that all the geometric forms have been accurately validated, including the symmetric axes.

![Figure 5: Symmetric geometric shapes along with their symmetric axes.](image)

4.5. Symmetric axes

By examining the angles at which the minimal mean directional error (MDE) occurs, the number of symmetric axes can be determined. Figure 6 displays photos of skin lesions with bilateral symmetry. The photos shown in Figure 7 exhibit skin lesions that possess a single axis of symmetry.

![Figure 6: Skin lesion images with bilateral symmetry.](image)
SAA: A novel skin lesion Shape Asymmetry Classification Analysis

Figure 7. Skin lesion images with one symmetric axis. (a) Skin lesion images. (b) one symmetric axis of skin lesion shape. (c) Symmetric axes induced skin lesion segmented image.

4.6. Classification result
The SVM model is used to categories skin lesion images into two categories: symmetric and asymmetric, using the SAA model. Figure 8 represents that SAA model has a high level of accuracy at 90%. Additionally, the model achieves favourable performance metrics such as F1-score, Precision and Recall, reaches to 0.89, 0.91 as well as 0.90 respectively. Furthermore, ROC curve for the model is 0.88 for both the classification labels.

![Confusion Matrix](image)

Figure 8. Results of SAA model. (a) Confusion matrix (b) Performance metrics (c) ROC curve

The SAA model is evaluated against established well known models such as Resnet_50 and AlexNet. The SAA model demonstrates higher performance across all performance parameters, including Accuracy, F1-score, precision, and recall, when compared to the current models. The performance measures across the various models are compared and has been shown in table 2. The SAA model focus on the shape feature alone to identify lesion asymmetry. The other deep learning models focuses on the other non-relevant features cause less accuracy in detection. Figure 9 illustrates the graphical depiction of the comparison between the model and the performance measures. Figure 10 illustrates that the variation in the gap between parallel lines does not have an impact on the determination of the mid-point, as well as the subsequent application of the RANSAC algorithm and calculation of MDE values.
Table 2. Comparison of performance metrics among the models.

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Accuracy</th>
<th>F1-score</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resnet_50</td>
<td>59</td>
<td>0.43</td>
<td>0.34</td>
<td>0.59</td>
</tr>
<tr>
<td>AlexNet</td>
<td>59</td>
<td>0.43</td>
<td>0.34</td>
<td>0.59</td>
</tr>
<tr>
<td><strong>Proposed model</strong></td>
<td><strong>90</strong></td>
<td><strong>0.89</strong></td>
<td><strong>0.91</strong></td>
<td><strong>0.90</strong></td>
</tr>
</tbody>
</table>

Figure 9. Result Analysis of SAA model with popular deep learning models. (a) Comparison of accuracy metric (b) Comparison of F1-score, recall and precision.

Figure 10. Effect on MDE for various Gaps taken while drawing parallel lines.

5. Conclusion and future work

A novel Shape Asymmetry Analysis model classifies skin lesion symmetry in dermoscopic images in this research. The PH2 dataset, which initially consisted of three labels, has been consolidated into two labels by unifying the categories of symmetric and symmetric with one axis into a single category called symmetric. Based on the findings, it can be concluded that the SAA model demonstrated more reliability and dominance. The top-ranked outcome achieved an exceptional accuracy of 90%, along with weighted F1-score of 0.89, precision of 0.91, recall values of 0.90. Additionally, the Receiver Operating Characteristic (ROC) yielded a value of 0.88, indicating a high level of precision and recall in determining symmetry. In addition to the categorization, the number of symmetric axes can be determined by examining the angles with minimal deviation error. When comparing the model to an efficient deep learning model on the PH2 dataset, it demonstrates higher performance across all performance criteria. Additionally, it was noted that the gap between the parallel lines does not have any impact on the computation of mean deviation error.

In future research, the integration of the output from the SAA model as a handcrafted feature, in combination with other non-handcrafted features extracted from skin lesion images, has shown promise in improving the accuracy of skin cancer identification. The SAA model demonstrates its applicability across a range of domains, including quality control in manufacturing, robotics and automation, geometry, and pattern recognition, among others. This feature allows for the evaluation of the symmetry of a shape and the calculation of symmetric angles, hence aiding in the creation of symmetrical lines.
References


