A Primary Morphological Classifier for Skin Lesion Images

Jules Matthew A. Macatangay
De La Salle University
2410 Taft Ave., Malate
Philippines 1004 Manila,
Metro Manila
jules_macatangay@dlsu.edu.ph

Conrado R. Ruiz, Jr.
De La Salle University
2410 Taft Ave., Malate
Philippines 1004 Manila,
Metro Manila
conrado.ruiz@dlsu.edu.ph

Richard P. Usatine
University of Texas
Health Science Center
8529 Raintree Woods Dr.
USA 78015 Fair Oaks
Ranch, Texas
usatine@uthscsa.edu

ABSTRACT

Classifying skin lesions, abnormal changes in skin, into their morphologies is the first step in diagnosing skin diseases. In dermatology, morphology is a categorization of a skin lesion's structure and appearance. Rather than directly classifying skin diseases, this research aims to explore classifying skin lesion images into primary morphologies. For preprocessing, k-means clustering for image segmentation and illumination equalization were applied. Additionally, features utilized considered color, texture, and shape. For classification, k-Nearest Neighbors, Decision Trees, Multilayer Perceptron, and Support Vector Machines were used. To evaluate the prototype, 10-fold cross validation was applied over a dataset assembled from online resources. In experimentation, the morphologies considered were macule, nodule, papule, and plaque. Moreover, different feature subsets were tested through feature selection experiments. Experimental results on the 4-class and 3-class tests show that of the classifiers selected, Decision Trees were best, having a Cohen's kappa of 0.503 and 0.558 respectively.

Keywords

Skin lesion, classification, machine learning, computer vision.

1 INTRODUCTION

Skin lesions are abnormalities in the surface of one's skin. These differences include changes in color, texture, and consistency. Various groups have become interested in applying technologies such as computer vision and machine learning techniques into the domain of skin lesions. Given that skin cancer is one of the most common, dangerous, and prevalent types of cancer, it has attracted many researchers to the development of effective automated detection of malignancies in skin lesions. With the technology available, many have done research on the computer-assisted analysis and diagnosis of skin lesions. For instance, [Oku13a], [Met14a], and [Sol16a] used images of skin lesions in determining whether the skin lesion is malignant or benign.

As much of the research focus on processing images of skin lesions has been on detecting malignancy, less have focused on classification besides the malignancy of a skin lesion. For one, [Ari12a] presented an automated dermatological diagnostic system that can classify skin lesion images into acne, eczema, psoriasis, tinea corporis, scabies, or vitiligo. On the other hand, [Yas14a] proposed the use of computer vision-based techniques to distinguish a skin lesion image as either acne, eczema, psoriasis, tinea corporis, scabies, vitiligo, foot ulcer, leprosy, or pityriasis rosea. It can be noted that one of the major differences between the research done by [Ari12a] and [Yas14a] is that the latter has a larger set of skin diseases as compared to the former. As much of the focus has been directly classifying from images to specific skin diseases, there is a research opportunity in shifting the focus of the classification problem.

Because of the vast amount of skin diseases that exist, another higher level categorization scheme that is inclusive of most, if not all, skin lesions may prove to be a useful and novel approach to the problem. Moreover, [Jam11a] wrote that the identification of a skin lesion's morphology is the first step in diagnosis. These along with the fact that morphology is indicative of a skin lesions' structure and appearance, makes morphology a proper means of categorizing skin lesions.

Morphology can be divided into primary and secondary. Primary morphology refers to the characteristic appearance of a skin lesion while secondary morphology refers to the temporal changes and modifications that occur on the skin lesion [Bol12a]. Only the primary morphologies will be explored in this research, as by

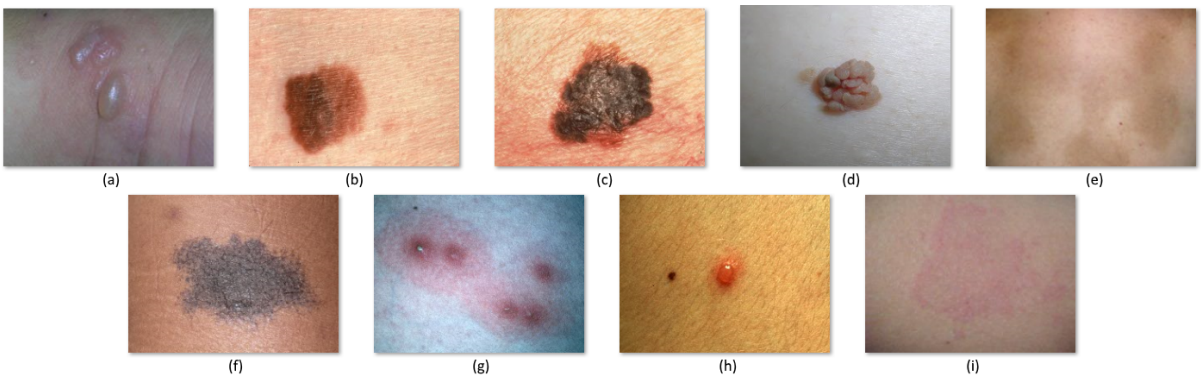

Figure 1: Examples of primary skin lesions: (a) Bulla (b) Macule (c) Nodule (d) Papule (e) Patch (f) Plaque, (g) Pustule, (h) Vesicle, and (i) Wheal.

definition they are more indicative of the skin lesion's core form and structure.

Thus, this research aims to explore classifying skin lesion images into primary morphologies. This is done by adapting techniques present in prior skin cancer malignancy and skin disease classification research to model a system that can classify images of skin lesions into the primary categorization of morphologies. As this research presents the novel research problem of classification by morphology, it sets the foundation for future work on further modeling of the diagnostic processes, or utilizing the higher level categorization to reduce complexity for classifying specific skin diseases.

2 RELATED WORK

Multiple research works have been done to create skin lesion classification systems. These works can be categorized, by their focus, into two groups. The first group, those focusing on skin malignancy classification, include studies determining whether a skin lesion is a malignant or benign, studies determining the severity of a skin cancer, and studies determining whether a skin cancer is a melanoma skin cancer or a non-melanoma skin cancer. The second group, those focusing on skin disease classification, include studies on classifying a skin lesion into a specific skin disease given a set of skin diseases. As of writing, there are no published studies focused on classifying skin lesions into their morphology.

2.1 Features and Feature Selection

There are a large variety of feature sets that have been defined in the literature. Most of the feature sets can be categorized into their focus. Primarily, features extracted from skin lesion images are those that describe the following elements of the skin lesion: color, texture, and shape.

Color is one common descriptor used in feature extraction for skin lesions. Commonly, an image can be expressed into a variety of color spaces. One example of this is to break down an image into the 3 color channels of the RGB color space. The resulting data can then be processed to extract features. In one study, [Yas14a] made use of the YCbCr color-encoding system in the extraction of color within a masked region.

Another crucial descriptor to skin lesions is their texture. Texture can be computationally expressed in a variety of ways [Mas13a]. One popular method is through the Grey Level Co-occurrence Matrix (GLCM). GLCM can be used to measure the spatial relationship between pixels, becoming a suitable means of extracting information on the texture of a skin lesion.

The shape of a skin lesion can be described in a variety of ways, such as area, size, and edge. [Yas14a] extracted features that quantify the area of a skin lesion and its apparent edges. [Yas14a] applied computations on the histogram and a masked region of the image for the area feature extraction, and applied the sobel operation on the masked region to extract the edges of the skin lesion.

In the study done by [Met14a], the Support Vector Machines Recursive Feature Elimination (SVMRFE) performed the best as compared to Information Gain and Correlation-based Feature Subset Selection. On the other hand, for [Mag09a], Correlation-based Feature Selection (CFS) performed the best as compared to using Principal Component Analysis (PCA) or Generalized Sequential Feature Selection (GSFS) for feature selection.

2.2 Classifiers

A lot of work has gone on the application of different classification methods in the context of skin lesions. A variety of methods can be applied to skin lesions: statistical classifiers, artificial neural networks, rule-based methods. [Yas14a] made use of a feed-forward back-propagation neural network in order to classify a skin lesion to one of 9 predefined skin diseases.

On the accuracy of such systems, many of the systems in the related works perform well in the detection and

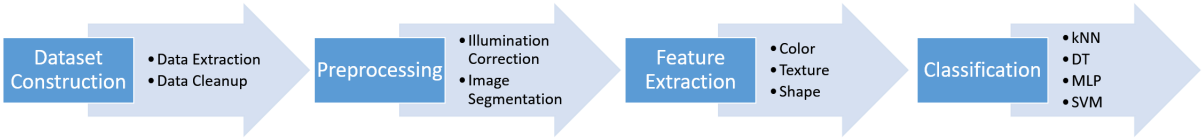

Figure 2: Process Flow

classification of skin cancer. For instance, [Yas14a] reports 91% - 97% accuracy for skin diseases that are low in elevation and 85% - 88% for skin diseases with high elevation. These studies served as the reference in selecting the various components of the classifier.

3 SKIN LESION MORPHOLOGIES

It is important to understand that morphology in this research is specifically in the context of dermatology. Morphology in dermatology is defined as the general appearance and structure of a particular skin lesion regardless of its function, etiology or pathophysiology [Bol12a]. Morphology can be further separated into primary and secondary morphology.

According to [Win86a], skin lesions can be grouped into two categories, primary and secondary morphologies. Primary morphologies differ in color or texture and are either acquired from birth, such as moles or birthmarks, or during a person's lifetime, as in the case of infectious diseases and allergic reactions. Secondary morphologies on the other hand are lesions that result from primary skin lesions, either as a natural progression or as a result of agitating the primary lesion. Because of this distinction, the categorization can be made to only consider primary morphologies.

Despite the vast amount of literature discussing skin lesions and morphology, different sources tend to list different sets of morphologies. For instance, the list presented by [Wol08a] is different from the one by [Pap04a], and by [Bic12a].

Regardless of these different listings, the description of each morphology is consistent amongst different references, so any of the references can be selected and utilized. For this study, a subset from the set of morphologies as listed by [Bic12a] and shown in Figure 1 is being used. This subset of morphologies was chosen to maximize the use of the data gathered for the research.

[Wel08a] provided the following brief description of the primary morphologies listed by [Bic12a]:

- Bulla - a fluid-filled circumscribed elevation of skin that is over 0.5 cm in diameter
- Macule - a small flat area with color or texture differing from surrounding skin
- Nodule - a solid mass in the skin that is palpated or elevated and is, in diameter of both width and depth, greater than 0.5 cm
- Papule - a solid elevation of skin that is less than 0.5 cm in diameter
- Patch - a large flat area with color or texture differing from surrounding skin
- Plaque - an elevated area of skin without substantial depth but is greater than 2 cm in diameter
- Pustule - an evident accumulation of pus in skin
- Vesicle - a fluid-filled circumscribed elevation of skin that is less than 0.5 cm in diameter
- Wheal - a white elevated compressible and faded area often surrounded by a red flare

As the descriptions for each morphology show, morphology in the dermatological sense is not only referring to shape but also referring to visual traits such as size, color, texture, and elevation. Furthermore, since no metric data will be utilized in this research, the system is limited in discriminating between morphologies. Another limitation is that this study focuses on skin lesions that are labelled as one morphology only, whereas cases wherein a particular skin lesion can be classified under multiple morphologies (e.g. a maculopapular rash belongs to both macule and papule) are not used in this research.

4 SKIN LESION CLASSIFIER

This section describes the concepts involved in the creation of a skin lesion classifier as outlined in Figure 2.

4.1 Preprocessing

Starting with an image, two major processes had to be made before proceeding with feature extraction: resolving nonuniform illumination and image segmentation.

Nonuniform illumination is the discrepancy between the images resulting from them being taken at different angles and lighting conditions, such as the example in Figure 3. For many instances, illumination normalization was enough to correct nonuniform illumination. However for other instances, Retinex [Zos13a] was used in tandem with segmentation. These methods simply refine the images and reduce nonuniform illumination rather than fix them completely.

Following the illumination refinements, the image will then be converted to the CIELab color space. After conversion, the k-means clustering algorithm was applied

on the a and b channels of the image in a CIELab color space to separate the image into two segments, as can be seen in Figure 4. Of the two segments, the one with the higher standard deviation was marked as the skin lesion, a step based on preliminary tests and observations. For a majority of cases, these steps were enough, whereas manual corrections had to be applied for the rest.

Once the two segments had been marked appropriately, the skin lesion segment was transformed into a mask. The mask was refined by filling holes, applying morphological opening, and removing small regions, a sample of which is at Figure 5. Lastly, this mask was then reapplied to the illumination refined image to get the properly divided skin lesion and healthy skin segments.

4.2 Features

After the image is properly segmented into normal skin and skin lesion, the segments undergo feature extraction. Many features extracted from skin lesion images covered in skin cancer and skin disease research revolve around three types: color, texture, or shape. Although many of these features were formulated for use in those specific research problems, there are features that could be adapted to classification by morphology.

4.2.1 Color

For color features, similar to [Sol16a], the common color features such as the mean and standard deviation of a value for each color channel in the RGB, HSV, and CIELab color spaces of the image were included. Additionally, by taking the mean color of the healthy skin, the difference between the mean color of the healthy skin and the mean color of the skin lesion can be computed and serve as a separate feature. These basic features, also used in this research, were previously utilized by works such as [Ari12a] and [Yas14a].

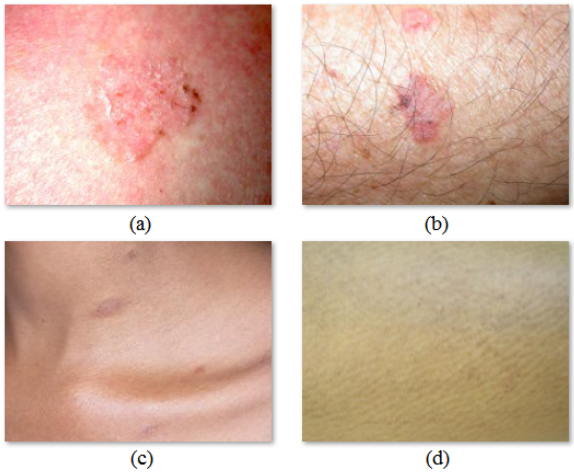

Figure 3: Issues in the dataset: (a) nonuniform illumination, (b) hair, (c) uneven skin surfaces, and (d) unclear skin lesion borders

Statistical measures such as uniformity or energy, and entropy applied to color channels also served as features, as done by [Met14a] for skin cancer classification. Both uniformity, computed through the equation

$$Uniformity = \sum_{k=0}^{M-1} p(k)^2 \quad (1)$$

and entropy, computed as

$$Entropy = \sum_{k=0}^{M-1} p(k)\log_2(p(k)) \quad (2)$$

define M as the number of bins of a distinct pixel value, and $p(k)$ as the probability associated with a specific color value k.

Moreover, the window-based color features of [Ari12a] were adapted to this research; in this case dropping the concentric circles scheme for a core, inner, and outer region division scheme.

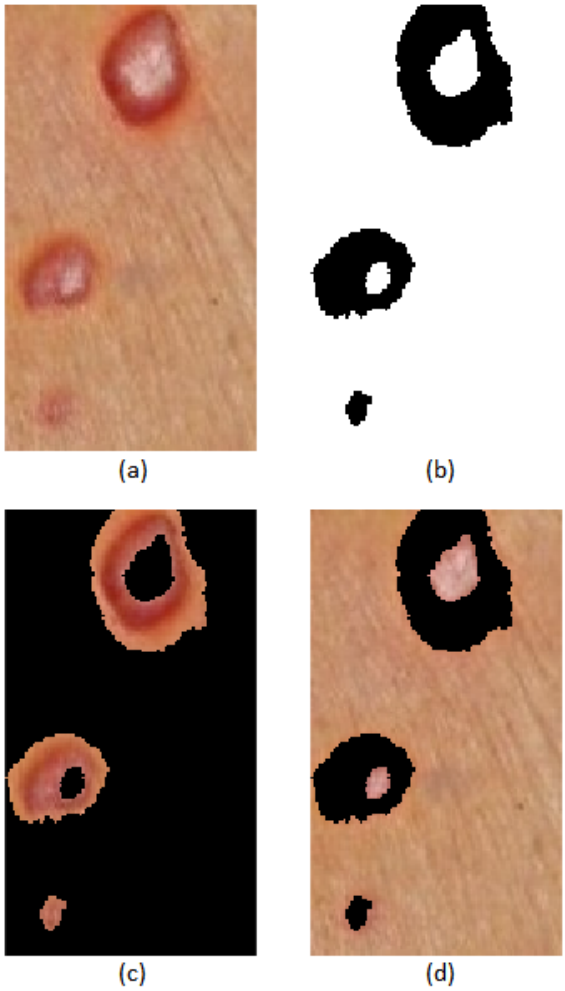

Figure 4: Segmentation steps: (a) Input, (b) K-means applied to a and b channels of image transformed to CIELab color space with k=2, (c) Skin Lesion, and (d) Healthy Skin

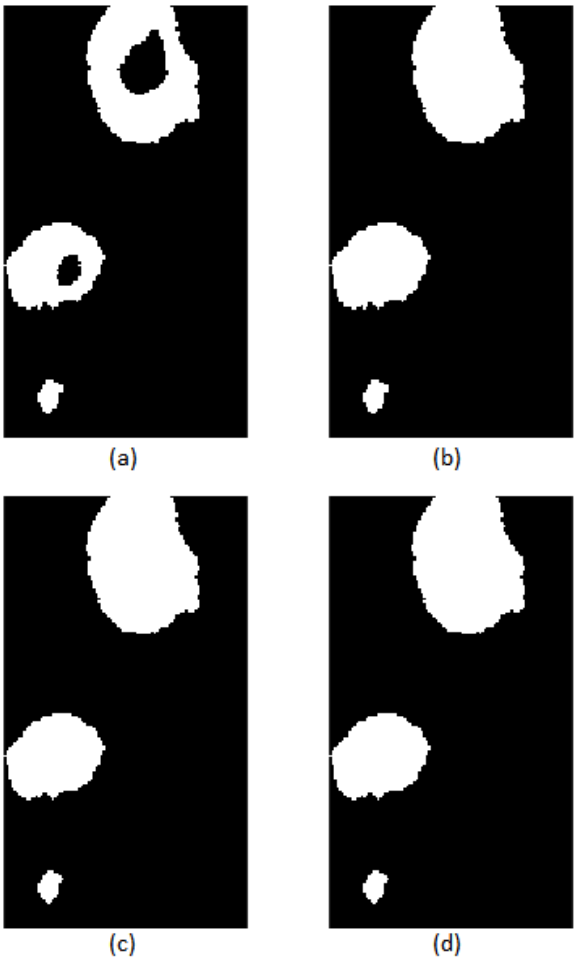

Figure 5: Mask refinements: (a) Input, (b) Filling holes, (c) Morphological opening, and (d) Removing small regions

Another feature, color asymmetry, was used by [Sma13a] for skin diseases through the chi-squared distance formula

$$D(h_1,h_2) = \sum_{i=1}^{N} \frac{(h_1-h_2)^2}{h_1+h_2} \tag{3}$$

wherein $D(h_1,h_2)$ is the chi-square distance of two histograms, N is the number of bins in histograms h_1 and h_2, and h_1 and h_2 are color histograms of opposite areas divided along a designated axis. This formula was applied over the two halves of an image as divided along either the minor or major axes; as was the feature designated for the relative distance of the intensity centroid and the mass centroid. In this research, the major and minor axes of a skin lesion is the same major and minor axes of an ellipse which has the same normalized second central moments as the skin lesion.

4.2.2 Texture

For texture features, many works have utilized various features available through the Gray-Level Co-occurence Matrix (GLCM). Generally, the common set which was selected for this research include energy, contrast, correlation and homogeneity. Energy is defined similar to how uniformity was treated in Equation 1, whereas the other three GLCM features are computed through the following equations

$$Contrast = \sum_{i=1,j=1}^{N} |i-j|^2 \, p(i,j) \tag{4}$$

$$Correlation = \sum_{i=1,j=1}^{N} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j} \tag{5}$$

$$Homogeneity = \sum_{i=1,j=1}^{N} \frac{p(i,j)}{1+|i-j|} \tag{6}$$

wherein i and j refer to the current column and row of the generated GLCM, μ is the mean of the current column or row, σ is the standard deviation of the current column or row, and $p(i,j)$ is the value found in the i-th column and j-th row of the GLCM.

Additionally, another set of features were derived from the Tamura texture features, which represent a visual description of texture based on human perception. As noted by [Cas02a], the first three features are enough for a metric on the general perception of texture. Moreover, as directionality is not visually emphasized in skin lesions and is more utilized in healthy skin, only contrast and coarseness were utilized for this research.

4.2.3 Shape

Shape features are not limited to the shape of the skin lesion, but also include how they are scattered and their number, as these are representative to how skin lesions present themselves. This means that the number of connected components serves as a feature as does distribution, as modified from [Ari12a] which is calculated through the formula

$$Distribution = \left(\frac{\sum_{i=1,j=1}^{N} dist_{ij}}{N} \right) \times \frac{1}{mmal} \tag{7}$$

wherein N is the number of regions, $dist_{ij}$ is the distance between the current region i and the current region j wherein both sets of regions refer to all skin lesions found, and *mmal* refers to the lowest measured minor axis length across all regions.

Moreover, [Ari12a] and [Paw14a] utilized features pertaining to the normalized area of the skin lesion, and the ratio of the min and max area, both of which were included for this research. The ratio of axis lengths and asymmetry along the major and minor axes as used by [Met14a] all served as features.

Data Source	Image Count
DermIS	6621
Global Skin Atlas	2551
DermAtlas	1114
Dermoscopy Atlas	394
Total	10680

Table 1: Raw dataset sources and image count

Morphology	Image Count
nodule	110
papule	108
plaque	93
macule	21
Total	332

Table 2: Final dataset size

For shape features focusing on the skin lesion border, features such as circularity, solidity, fractal dimension index and border solidity all served as varying means of measuring how defined, ragged, and uniform the borders appear to be [Met14a].

4.3 Feature Selection

Due to the large amount of features available, feature selection was also explored through two methods. Primarily, a reduced feature set was assembled based on preliminary tests and observations.

Alternatively, a genetic algorithm was also used to come up with a different reduced feature set. Genetic algorithms are based on the process of natural evolution. Generally, genetic algorithms use a heuristic to generate a solution to a specified problem.

4.4 Classification Methods

After feature extraction, different classifiers were then trained using a fraction of the generated dataset. Two of the classifiers chosen are standard algorithms utilized in machine learning, while the other two are based on what skin cancer and skin disease classification literature support.

4.4.1 K-Nearest Neighbors

The first classifier, K-Nearest Neighbor (kNN), is an algorithm that compares a given test sample described by a number of attributes to samples with similar attributes. Its name is derived from the fact that given a new sample, the algorithm searches the n-dimensional pattern space for k number of training samples similar to it based on a given distance metric. This algorithm is one of the simplest of the machine learning algorithms and yet still performs relatively well in particular classification problems.

4.4.2 Decision Trees

The second classifier, Decision Trees (DT), are inverted tree-like graphs that represents a prediction model for a target attribute given several input attributes. Each interior node of the tree corresponds to an input attribute, while the leaf nodes represents the predicted value of the target attribute upon traversal of the tree. Algorithms such as ID3 and C4.5 create decision trees based on attribute values of a dataset. In particular, these algorithms utilize information gain in order to designate the root node and recursively partition the values of the attributes into different nodes. It should also be noted that there is a chance that the algorithm may not be able to accommodate all data entries from the training data due to noise.

4.4.3 Multilayer Perceptron

According to [Ari12a], [Paw14a], and [Yas14a], Feed-Forward Back Propagation Artificial Neural Networks perform the best in classification problems on skin diseases. Artificial neural networks (ANN) consist of smaller processing units or neurons that are highly interconnected to form a computational model. Training involves having each neuron adjust their weights to accommodate the current input. Furthermore, ANN are popular in dermoscopic image analysis according to both [Mag09a] and [Mas13a]. A Multilayer Perceptron (MLP) is a type of artificial neural network that is an extension of perceptrons, and is said to be the simplest kind of feed-forward neural network.

4.4.4 Support Vector Machines

According to [Mag09a] and [Mas13a], Support Vector Machines (SVM) perform the best in classification problems concerning skin lesion malignancy. SVMs work based on statistical learning theory, essentially finding the optimal hyperplane between classes in a dataset through the resolution of an optimization problem.

5 EXPERIMENTS AND RESULTS

Given the flow outlined in Figure 2 and discussed in Section 3, specific experiments were designed and ran to evaluate the classifiers generated from the processes described.

5.1 Dataset and Experiment Setup

First, images and textual data were gathered from 4 data sources: DermIS [Hei03a], Global Skin Atlas [Aus05a], the Interactive Dermatology Atlas [Usa06a], and Dermoscopy Atlas [Aus07a]. After extraction, the data was combined based on the morphologies tagged and filtered further to images that were tagged with only one morphology belonging to the four utilized in this research. Furthermore, instances were also excluded from processing due to various issues such as uneven skin surfaces and unclear skin lesion borders,

Classifier	w/ Feat Select	accuracy	kappa
kNN	No	46.06%	0.215
	Yes	49.40%	0.264
DT	No	60.38%	0.427
	Yes	65.48%	0.501
MLP	No	67.17%	0.530
	Yes	65.14%	0.502

Table 3: Accuracy and Cohen's kappa for 4 Class dataset with or without genetic feature selection applied

Classifier	Feature Set	accuracy	kappa
KNN	Base	46.06%	0.215
	Reduced	56.10%	0.363
DT	Base	60.38%	0.427
	Reduced	65.63%	0.503
MLP	Base	67.17%	0.530
	Reduced	64.03%	0.486
SVM	Base	32.16%	-0.014
	Reduced	59.57%	0.408

Table 4: Accuracy and Cohen's kappa for 4 Class dataset with differing feature sets

such as those in Figure 3. Due to the aforementioned constraints, the number of instances in the dataset dwindled from 10,680 in Table 1 to 332 in Table 2.

For the configurations of the classifiers, preliminary tests were done to find the configurations that performed well for the classification problem. Firstly, the kNN classifier was set to use a weighted voting scheme with $k = 5$. For the DT, the criterion of accuracy was utilized and the classifier was set to have a maximal depth of 20, with pruning and pre-pruning active. The MLP utilized was set to have 10 maximum training cycles and run over 10 generations with 4 MLPs per ensemble. Lastly, A C-SVC type SVM with the rbf kernel was utilized.

5.2 Feature Selection Experiments

For this test, Table 3 shows the result of the genetic feature selection algorithm as trained using the SVM classifier. Alternatively, Table 4 shows the comparison between utilizing the full feature set or a manually reduced feature set. The reduction was done by removing features that correspond to any of the HSV and CIELab color space channels, leaving just color features based on intensity.

As can be observed from both Table 3 and 4, applying the genetic feature selection to the feature set did not affect the performance of the classifiers consistently. Moreover, the kNN classifier performed less as compared to both DT and MLP, which may be attributed to the still high dimensionality of the feature space even after many features had been filtered out.

On the contrary, utilizing the reduced feature set yielded similar if not better results than utilizing the whole feature set, with MLP being the exception. Furthermore, as the reduction on the feature set is high for this test, the improvements for the kNN and SVM classifiers is easily evident. As utilizing the reduced feature set allows for faster processing and utilizes a smaller feature set, it was the feature set utilized for the succeeding 4 class and 3 class tests.

An important insight is that although the reduced feature set provided better improvements than genetic feature selection, as Table 5 includes features relating to the removed HSV and CIELab colors pace channel features, these features not present in the reduced feature set may still hold considerable value.

5.3 4 Class

The 4 class test were done over the four specific classes, namely: macule, nodule, papule, and plaque.

As can be observed from Table 6, nodules and papules were the two classes that the system had the easiest time distinguishing. All of the classifiers seem to have difficulty with detecting skin lesions that are macules but did better with plaques. Given that the data is unbalanced in that there are fewer instances of macule than the other classes, this may have skewed the result against the class. For this test, DT and MLP performed better as compared to kNN and SVM.

5.4 3 Class

The 3 class tests were those that included only three classes: nodule, papule, and plaque. For this test, macule was removed as it contained a significantly fewer amount of cases as compared to the other three classes.

In Table 7, as compared to that of Table 6, the removal of the macule class showed improvements but certain trends remain. For instance, the classes of nodule and papule have higher f-measures than plaque which is still lagging behind. This may be attributed to the fact that papule and plaque are very similar. For instance, Figure 7 shows a set of plaque skin lesions tagged correctly and incorrectly which visually look to have the same texture and border definition. However, possibly due to the differences in their distribution and color, the samples on the left were misclassified.

Additionally, it must be noted that all three classes in the 3 class test are all categorically classified as raised

Distribution	Color(Inner - H)
Tamura Contrast	Color(Outer - I)
Contrast(Lesion)	Mean a(Lesion-Skin)
Uniformity b	Solidity
Circularity	Correlation(Lesion/Skin)
Mean b(Lesion-Skin)	Color(Inner - I)
Relative a Centroid	Color P. Asym.(Minor - I)

Table 5: Features in pruned decision tree based from genetic feature selection feature set

	macule	nodule	papule	plaque
kNN	0.00%	64.67%	69.54%	35.48%
DT	29.03%	71.77%	72.38%	55.11%
MLP	28.95%	72.22%	71.22%	52.91%
SVM	0.00%	64.74%	72.48%	43.83%

Table 6: F-measure for 4 Class test

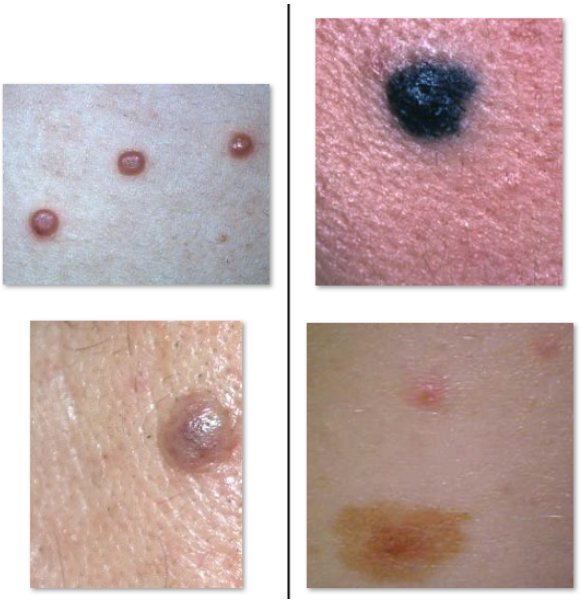

Figure 6: Sample images: (left) papules classified as nodules; (right) papules classified as papules

skin lesions (with macule being part of the flat morphologies) and so share many similarities with one another, making discrimination between them more difficult. For instance, Figures 6, 7, and 8 all show a series of nodules, papules, and plaques, all of which appear to be above the general elevation of skin. Based on the results shown in Table 7 for this test, DT performed the best, only slightly falling behind MLP in distinguishing nodules. Given Table 8, it can be inferred that based on both accuracy and Cohen's Kappa, the Decision Tree classifier is able to create the best classification model, a notion supported by all tests recorded.

	nodule	papule	plaque
kNN	68.58%	70.40%	43.12%
DT	77.80%	71.26%	60.50%
MLP	77.94%	70.87%	56.98%
SVM	69.48%	70.84%	42.27%

Table 7: F-measure for 3 Class test

	4 Classes		3 Classes	
	accuracy	kappa	accuracy	kappa
kNN	56.10%	0.363	62.23%	0.429
DT	65.63%	0.503	70.71%	0.558
MLP	64.03%	0.486	69.39%	0.539
SVM	59.57%	0.408	62.55%	0.434

Table 8: Accuracy and Cohen's kappa for 4 Class and 3 Class tests

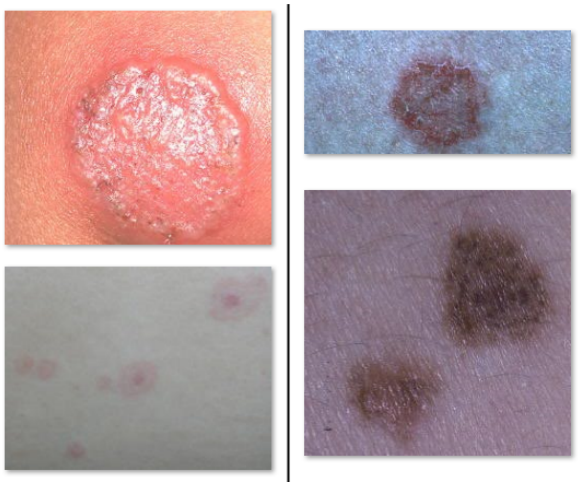

Figure 7: Sample images: (left) plaques classified as papules; (right) plaques classified as plaques

6 CONCLUSION

This research study was able to explore classifying skin lesion images into primary morphologies. Out of the four classifiers, Decision Trees performed best and thus is the recommended classifier both due to its resiliency and performance. For feature selection, genetic feature selection provides inconsistent results, and utilizing a reduced feature set showed an increase in performance in all except for MLP. In conclusion, the classification of skin lesion images by morphology is possible. However, more research is needed for significant use of a system based on the technology can be utilized, especially as the research area has not been thoroughly explored.

Further research can benefit from a larger and more consistent dataset, especially as the current dataset has many variations in lighting, scaling, and resolution. A new dataset built through strict guidelines may provide valuable insight into the deeper exploration of classification by morphology. Additionally, testing a new feature set, formulating new features specific to the research problem, and exploring features outside prior skin cancer and skin disease classification works may prove beneficial. Moreover, a multi-tier classification scheme may be possible given that morphologies can be grouped (e.g. papule, nodule, and plaque are all raised)

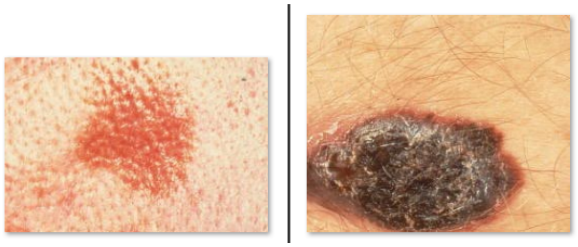

Figure 8: Sample images: (left) nodule classified as plaque; (right) nodule classified as nodule

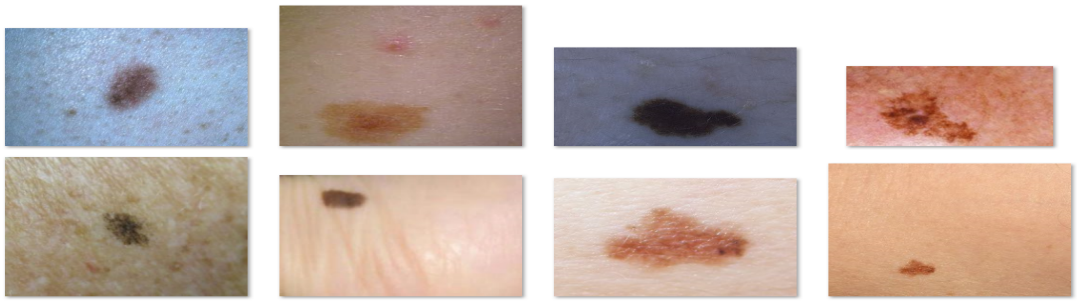

Figure 9: Sample macules correctly classified

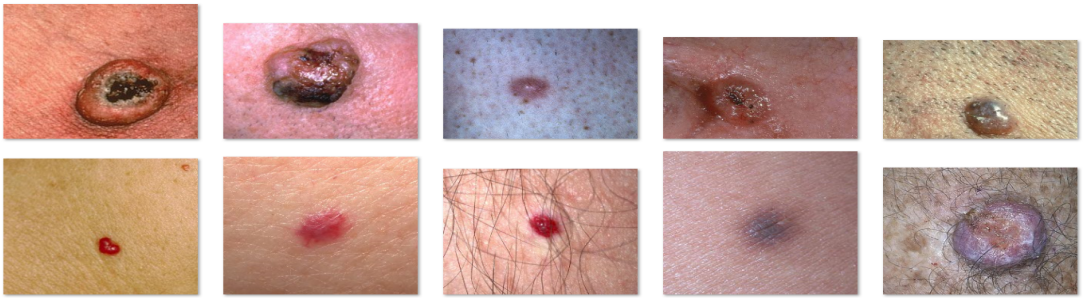

Figure 10: Sample nodules correctly classified

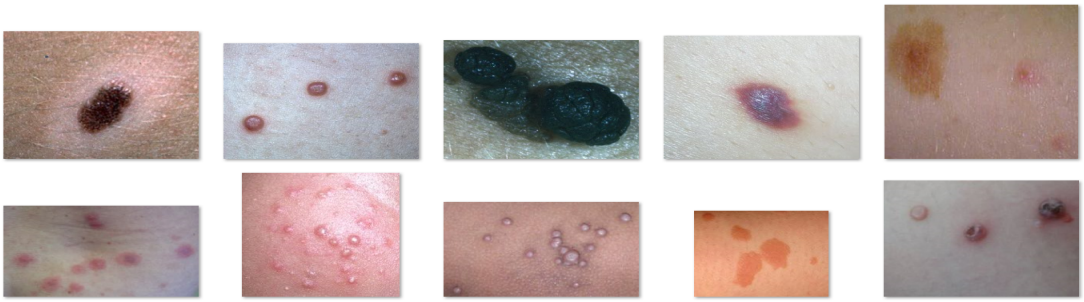

Figure 11: Sample papules correctly classified

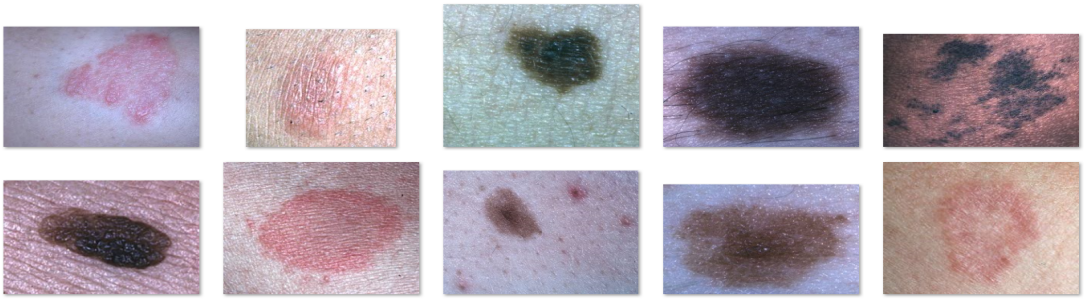

Figure 12: Sample plaques correctly classified

and that some pairs of morphologies are very similar to each other (e.g. papule and plaque are similar but differing in size). Also, further research can be made on multilabel classification to accommodate skin lesions that fall on multiple morphologies. Alternatively, the use of convolutional neural networks (CNNs) with regards to the research problem is also worth investigating. Lastly, further work can be on measuring the benefits of including the automatically tagged morphologies as a component in the classification of skin diseases.

7 ACKNOWLEDGMENTS

This research was made possible with funding from the University Research Coordination Office (URCO) of the De La Salle University. Many thanks to Dr. Arnulfo Azcarraga, Dr. Joel Ilao, Dr. Maria Franchesca Quinio and Dr. Erin Jane Tababa for their input, and to DermIS [Hei03a], Global Skin Atlas [Aus05a], the Interactive Dermatology Atlas [Usa06a], and Dermoscopy Atlas [Aus07a] for the dataset.

8 REFERENCES

[Ari12a] Arifin, M. S., Kibria, M. G., Firoze, A., Amini, M. A., & Yan, H. (2012). Dermatological disease diagnosis using color-skin images. In International Conference on Machine Learning and Cybernetics (ICMLC 2012), 5, 1675-1680.

[Aus05a] The Skin Cancer Society of Australia. (2005). Global Skin Atlas. Retrieved from http://www.globalskinatlas.com/index.cfm

[Aus07a] The Skin Cancer Society of Australia. (2007). Dermoscopy Atlas. Retrieved from http://www.dermoscopyatlas.com/index.cfm

[Bic12a] Bickley, L., & Szilagyi, P. G. (2012). Bates' guide to physical examination and history-taking. Philadelphia, PA: Lippincott Williams & Wilkins.

[Bol12a] Bolognia, J. L., Jorizzo, J. L., Schaffer, J. V., Cerroni, L., Heymann, W. R., & Callen, J. P. (2012). Dermatology (Vol. 2). New York, NY: Mosby.

[Cas02a] Castelli, V., & Bergman, L. D. (2002). Image Databases: Search and Retrieval of Digital Imagery. New York, NY: John Wiley & Sons.

[Hei03a] University of Heidelberg - Department of Clinical Social Medicine, & University of Erlangen - Department of Dermatology. (2003). DermIS - Dermatology Information System. Retrieved from http://www.dermis.net/

[Jam11a] James, W. D., Elston, D., & Berger, T. (2011). Andrew's diseases of the skin: clinical dermatology (11th ed.). London, UK: Saunders Elsevier .

[Mag09a] Maglogiannis, I., & Doukas, C. N. (2009). Overview of advanced computer vision systems for skin lesions characterization. IEEE transactions on information technology in biomedicine, 13(5), 721-733.

[Mas13a] Masood, A., & Ali Al-Jumaily, A. (2013). Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms. International journal of biomedical imaging, 2013, 22.

[Met14a] Mete, M., & Sirakov, N. M. (2014). Optimal set of features for accurate skin cancer diagnosis. In 2014 IEEE International Conference on Image Processing (ICIP), 2256-2260.

[Oku13a] Okuboyejo, D., Olugbara, O., & Odunaike, S. (2013). Automating skin disease diagnosis using image classification. In Proceedings of the World Congress on Engineering and Computer Science, 2.

[Pap04a] Papier, A., Chalmers, R. J., Byrnes, J. A., & Goldsmith, L. A. (2004). Framework for improved communication: the Dermatology Lexicon Project. Journal of the American Academy of Dermatology, 50(4), 630-634.

[Paw14a] Pawar, M., Sharma, D. K., & Giri, R. N. (2014). Multiclass Skin Disease Classification Using Neural Network. International Journal of Computer Science and Information Technology Research. 2, 189-193.

[Sma13a] Smaoui, N., & Bessassi, S. (2013). A developed system for melanoma diagnosis. International Journal of Computer Vision and Signal Processing, 3(1), 10-17.

[Sol16a] Solomon, A. M., Murali, A., Sruthi, R. B., Sreekavya, M. K., Sasidharan, S., & Thomas, L. (2016). Identification of Skin Cancer based on Colour, Subregion and Texture. International Journal of Engineering Science, 8331

[Usa06a] Usatine, P., & Madden, B. (2006). DermAtlas - The Interactive Dermatology Atlas. Retrieved from http://www.dermatlas.net/index.cfm

[Wel08a] Weller, R., Hunter, J., Savin, J., & Dahl, M. (2008). Clinical Dermatology (4th ed.). Massachusetts, MA: Malden Publishing.

[Win86a] Winkelmann, R. K. (1986). Glossary of basic dermatology lesions. The International League of Dermatological Societies Committee on Nomenclature. Acta dermato-venereologica. Supplementum, 130, 1-16.

[Wol08a] Wolf, K., Goldsmith, L., Katz, S. I., Gilchrest, B., Paller, A. S., & Leffell, D. J. (2008). Fitzpatrick's Dermatology in General Medicine (7th ed.). USA: McGraw-Hill.

[Yas14a] Yasir, R., Rahman, M. A., & Ahmed, N. (2014). Dermatological disease detection using image processing and artificial neural network. In 2014 International Conference on Electrical and Computer Engineering (ICECE), 687-690.

[Zos13a] Zosso, D., Tran, G., & Osher, S. (2013). A unifying retinex model based on non-local differential operators. In International Society for Optics and Photonics IS&T/SPIE Electronic Imaging, 865702.