INCITING EXPLICIT FEATURES USING GENETIC PROGRAMMING FOR MELANOMA DETECTION

[1] R. Dhivya Raj Kumari, [2] M.Tamilselvi, [3] K.Kapilavani, [4] S.Bhuvaneswari,
[1,2] Students, [3,4] Faculty,
[1,2,3,4] Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar

ABSTRACT

Melanoma is that the deadliest kind of skin cancer that causes around 75% of deaths worldwide. However, most of the skin cancers will be cured, especially if detected and treated early.Existing approaches have employed various feature extraction methods, where differing types of features are used individually for skin image classification which cannot provide sufficient information to the classification algorithm necessary to discriminate between classes, resulting in sub-optimal performance.This study develops a completely unique skin image classification method using multi-tree genetic programming (GP).

To capture local information from gray and color skin images, Local Binary Pattern is employed during this work.In addition, for capturing global information, variation in color within the lesion and therefore the skin regions, and domain-specific lesion border shape features are extracted. GP with a multi-tree representation is utilized to use multiple kinds of features. Genetic operators like crossover and mutation are designed accordingly so as to pick out one form of features at terminals in one tree of the GP individual.

The performance of the proposed method is assessed using two skin image datasets having images captured from multiple modalities, and compared with six most typically used classification algorithms additionally because the standard (single-tree) wrapper and embedded GP methods.The results show that the proposed method has significantly outperformed of these classification methods. Being interpretable and fast in terms of the computation time, this method can help dermatologist identify prominent skinimage.

***Keywords*- Genetic programming, Feature selection, Feature construction, Image classification, Melanoma detection.**

1. INTRODUCTION

Cancer may be a deadly disease within which an abnormal growth of cells tends to proliferate in an untrammeled way that ends up in formation of malignant lumps. Melanoma is one among the fatal style of cancer, accounting for nearly 40% of occurrences worldwide. quite 2 people die of skin cancer within the U.S every hour. However early diagnosis of skin cancer is extremely curable with a survival rate of nearly 92%.One of the dermatologists most well liked imaging techniques is dermoscopy , whichquantify characteristics of Asymmetry , Border, Color, and Dermoscopic structure. It helps to effectively segregate different sorts of skin images.

Researchers are focused to plot methods that integrate both local and global features extracted from skin cancer images. The study mainly focused on the classical Machine Learning work flow that consists of pre-processing,segmentation, and classification. The Multi-tree Genetic Programming utilizes its implicit feature selection property to automatically select the relevant features as its terminal.

The proposed method enhances the accuracy of pre-processing techniques without losing informative features. For effective results a sturdy skin cancer categorized method is expanded over multiple image modalities. The Segmentation process clearly distinguish the skin cancer from benign cells. The Binary Mask is additionally incorporated to the classification.

2. RELATED WORK

GP has been widely explored for image analysis. A group of requirements for fitness function, terminal set, and performance set in GP has been outlined necessary to get effective optimal filters in X-ray coronarograms and brain MRI. . Their method works by evolving multiple trees in an exceedingly GP individual on the training data, where each tree operates as a binary classifier. The tree producing the most effective performance on the training data is employed to check the performance on the test data. they need identified the important image features by analyzing the great evolved GP programs, which might help dermatologists to create diagnosis in real world situations.

Some existing approaches developed classification methods for melanoma detection where various features are extracted from skin images. These methods assessed the goodness of those features individually using different machine learning classification algorithms. However, they lack employing a combination of various forms of features concurrently to realize performance gains. Performance is improve by utilizing of these features concurrently by designing a good way of mixing these differing kinds of features.

Most existing methods have used only 1 image modality (images captured from one instrument) to check the performance of their method(s).

However, in real-world situations, there are images captured from different instruments and hence, these methods, developed for one image modality, can't be applied to or may perform poorly on other image modalities. Hence, there's a necessity for a classification method for skin images which, having sufficient informative features, has the flexibility to be applied to multiple image modalities, easily interpretable so as to guide the dermatologist, and ready to discriminate between various classes of skin cancers.

3. PROBLEM DEFINITION

Skin is that the outermost layer of human body and it acts as a protective covering from external agents. . Skin acts as a primary layer of defensive system against germs and other foreign bodies. There are many diseases affecting skin. Among that, most dangerous and life-threatening condition is skin cancer. It could be a deadly condition affecting the skin. Late diagnosis and lack of effective treatment are the most reasons. skin cancer, if not treated properly can result in death of the patient.

3.1 SYSTEM MODEL

The proposed method operates on a group of predefined/extracted features which include local and global information about the skin images. The local features are extracted with the assistance of LBP descriptor which works with the pixel values and may significantly capture informative features about various skin properties like lines/streaks, blobs, homogeneous regions, and irregular border patterns. the worldwide features are extracted by specializing in shape and color variation characteristics of skin lesions.

These features are of utmost importance because without using these human crafted features, it's difficult to realize good performance for such a difficult task as skin image classification. These global features capture the properties of asymmetry, border, color and diameter (ABCD) rule of dermoscopy, which plays an important role for the dermatologist in distinguishing malignant from benign images. Hence, incorporating these informative features help the classifier learn better and produce a good model.

This section provides an in-depth description of the proposed MTGP wrapper method, which starts by presenting an outline of the algorithm to evolve a GP individual so as to focus on the key components of our proposed method, and the way the constructed features from the evolved individual are used for classification.

Then the program structure, i.e., the terminal and therefore the function sets, the crossover and mutation operators, and therefore the fitness function, are discussed. The melanoma detection provides the subsequent set of modules:

- Pre-Processing
- Segmentation
- Classification

3.1.1 PRE-PROCESSING

Image processing is any variety of signal processing that the input is a picture, like a photograph or video frame; the output of image processing is also either a picture or a collection of characteristics or parameters associated with the image. Image processing operations may be roughly divided into three major categories, compression, Image Enhancement and Restoration and Measurement Extraction.

3.1.2. IMAGE ENHANCEMENT

Image enhancement is that the improvement of digital image quality (wanted e.g. for visual inspection or for machine analysis), without knowledge about the source of degradation.

If the source of degradation is thought, one calls the method image restoration.Frequently space-invariant grey value transformations also are in serious trouble contrast stretching, range compression, etc. The critical distribution is that the ratio of every grey value, the greyvalue histogram.

3.1.3. RGB TO GREYSCALE CONVERSION

Conversion of a color image to grayscale isn't unique; different weighting of the colour channels effectively represent the effect of shooting black-and-white film with different-colored photographic filters on the cameras.

To convert a color from a color space supported an RGB color model to a grayscale representation of its luminance, weighted sums must be calculated in a very linear RGB space, that is, after the gamma compression function has been removed first via gamma expansion. I = rgb2gray(RGB) converts the TrueColor image RGB to the grayscale intensity image I. The rgb2gray function converts RGB images to grayscale by eliminating the hue and saturation information while retaining the luminance.

3.1.4. MEDIAN FILTER

International Journal of Advanced Development in Science and Technology
Volume : 3 Issue : 03 April 2021 www.ijadst.com e-ISSN: 2582-1059

Median filtering may be a nonlinear process useful in reducing impulsive, or salt-and-pepper noise. it's also useful in preserving edges in a picture while reducing random noise. The median filter considers each pixel within the image successively and appears at its nearby neighbors to determine whether or not it's representative of its surroundings. rather than simply replacing the pixel value with the mean of neighboring pixel values, it replaces it with the median of these values.

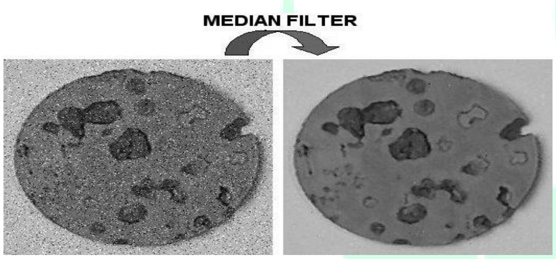

Fig 1: Median Filter

The median is calculated by first sorting all the pixel values from the encircling neighborhood into numerical order so replacing the pixel being considered with the center pixel value.

3.1.5. DIGITAL DATA COMPRESSION

Many image file formats use data compression to cut back file size and save space for storing. Digital compression of images may happen within the camera, or is drained the pc with the image editor. When images are stored in JPEG format, compression has already taken place. Both cameras and computer programs allow the user to line the amount of compression.

3.1.6. IMAGE RESTORATION

The aim of image restoration is to "compensate for" or "undo" defects which degrade a picture. Degradation comes in many forms like motion blur, noise, and camera misfocus.

3.1.7. SEGMENTATION

Image segmentation is that the process of partitioning a digital image into multiple segments (sets of pixels, also referred to as super-pixels). The goal of segmentation is to simplify and/or change the representation of a picture into something that's more meaningful and easier to research. Image segmentation is often accustomed locate objects and bounds (lines, curves, etc.) in images.

3.1.8. GRAY-LEVEL THRESHOLDING

This technique is an example of the regional approach, which suggests grouping pixels into distinct regions or objects. When using this method for image segmentation, all pixels that have gray levels up to or above a particular threshold are assigned to the item. the remainder of the pixels are assigned to the object's background. If the image contains over one object, then more threshold levels are often accustomed do the segmentation.

3.1.9. GRADIENT-BASED SEGMENTATION

This technique concentrates on the boundaries between different regions. It describes three techniques that are supported the gradient technique. they're the boundary tracking technique, the gradient image thresholding, and Laplacian edge detection. We describe each of those techniques briefly.

3.1.10. LAPLACIAN EDGE DETECTION

The Laplacian is a scalar, second-order derivative operator that, for a two-dimensional function f(x, y), can be given by

The Laplacian will produce an abrupt zero-crossing at an edge. If a noise-free image has sharp edges, the Laplacian can find them. In the presence of noise, however, low-pass filtering must be used prior to using the Laplacian.

3.1.11. CLASSIFICATION

A classification algorithm (such as a decision tree) is trained on the transformed training set. The learned classifier is then applied to the transformed test set to obtain the final test classification performance.

Fitness Function: The balanced classification accuracy is used as the fitness function, which is defined as

$$\text{Fitness} = \frac{1}{m}\sum_{i=1}^{m}\frac{TP_i}{TP_i + FN_i}$$

where m is the number of classes, T P refers to the true positive, F N refers to the false negative, and the ratio $\frac{TP_i}{TP_i + FN_i}$ represents the true positive rate of a class.

Crossover and Mutation To meet the objective of having only one type of features in a single GP tree, genetic operators, such as crossover and mutation, are designed accordingly, which is called same-index crossover/mutation. The step-by-step process is given in Algorithms 1 and 2. This crossover/mutation guarantees that the GP individual evolved at the end of the evolutionary process, consists of four trees where each tree evolves from a single type of features.

Algorithm 1: Same-Index Crossover.

1: function CROSSOVER P1, P2 Two GP Individuals (parents), each having four trees

2: for i = 1 to 4 do

3: XOVER(P1 i , P2 i) Crossover between trees having same

4: type of features as terminals

5: end for

6: return C1, C2 The two children obtained after XOVER

7: end function

Algorithm 2: Same-Index Mutation.

1: function MUTATION(P1) One GP Individual (parent) having four trees

2: for i = 1 to 4 do

3: P1 ← init (Ti) Generate a new tree with

4: a single type of features

5: MUTATE(Pi, P1) Mutate the tree from parent

6: individual with the new generated tree,

7: both having the same type of features

8: end for

9: returnC1 One child obtained after MUTATE

10: end function

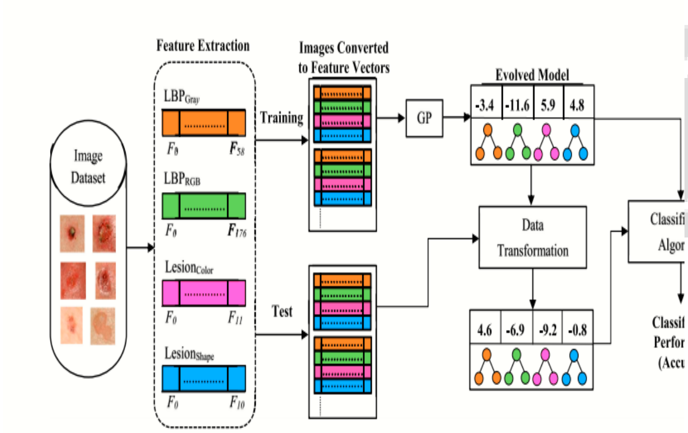

Fig 2:System Architecture Diagram

Overview of the proposed method: Each image in the dataset is input to four feature extraction methods to obtain four feature vectors, namely LBPGray, LBPRGB LesionColor, and LesionShape, for each image.

The training set is given to GP to evolve four trees each based on a single feature vector.

Using these four trees (constructed features), the training and test sets are transformed into new training and test sets. The transformed training set is provided to a classification algorithm (such as a decision tree) to evolve a classification model.

The learned classification model is applied to the transformed test set to obtain the test classification performance.

4. RESULTS AND DISCUSSION

The image acquired is then pre-processed to extract various local and global features. The pre-processing involves image acquisition, noise removal, and image enhancement processes.

Various datasets were collected and one example among the collected dataset are preprocessed to form an equal aspects ratio so that it can be made ready for training with the model.

International Journal of Advanced Development in Science and Technology

Volume : 3 Issue : 03 April 2021 www.ijadst.com e-ISSN: 2582-1059

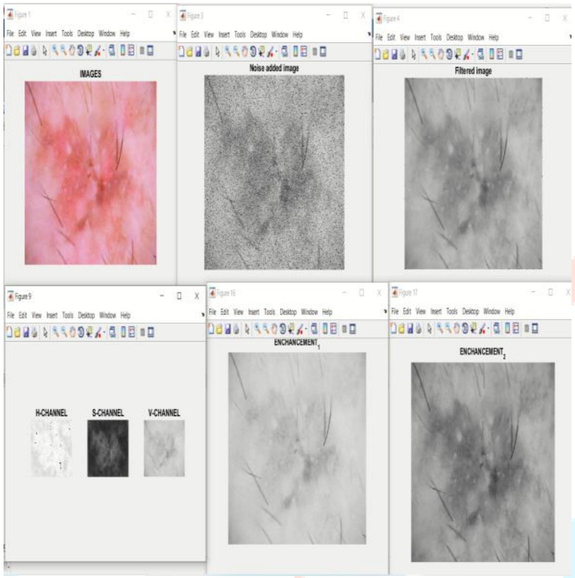

Fig 3: Pre-processing

After processing the image, the cancer affected part is segmented in the segmentation process. For segmentation it takes about 900 to 1500 iterations.

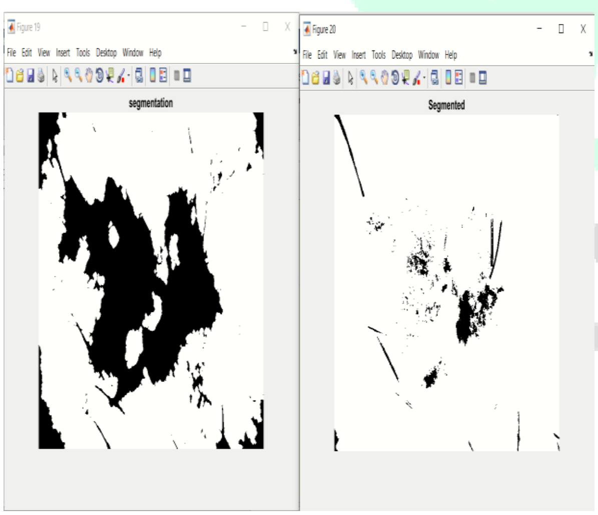

Fig 4: Segmentation

Finally the image is classified and the stage to the cancer gets affected is determined.

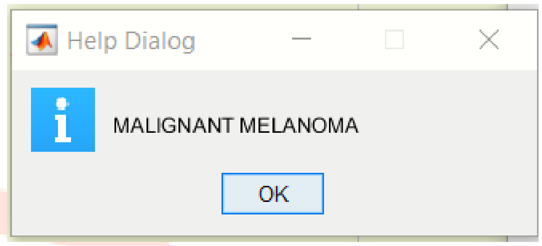

Fig 5: Help Dialog

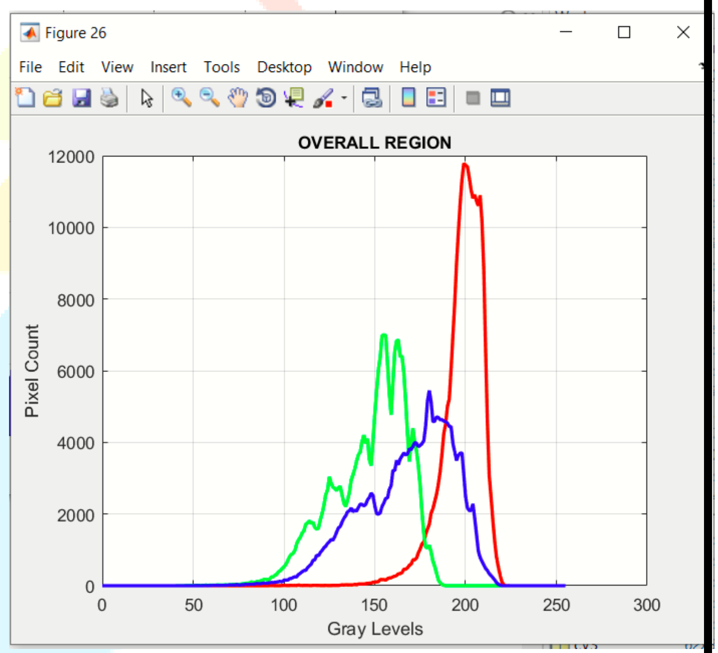

Fig 6 Classification

5.CONCLUSION

In this work, a completely unique method for skin cancer image classification using MTGP in a very wrapper approach has been developed. The proposed method utilizes various local and global features extracted from skin cancer images. These features have sufficient information associated with pixel-based RGB and gray-level

International Journal of Advanced Development in Science and Technology
IJADST Volume : 3 Issue : 03 April 2021 www.ijadst.com e-ISSN: 2582-1059

properties, domain-specific geometrical shape characteristics, and variation in color within the lesion and also the skin areas.These type of pre-extracted features are given to multi-tree GP to come up with four trees in a very single GP individual by designing suitable genetic operators like same-index-crossover/mutation. This kind of crossover/mutation guarantees that each tree evolves from a single type of features. We have also analyzed an interesting behavior to pick an acceptable feature extraction method so as to classify well a selected type of images taken from a particular optical instrument. We found that the local pixel-based features provide good discriminating knowledge to classify specialized (dermoscopy) images. On the other hand, global color variation and border shape features have more potential to discriminate images captured from a typical camera.

6. REFERENCES

[1] R. Kasmi and K. Mokrani, " Classification of malignant melanoma and benign skin lesions : Implementation of automatic ABCD rule," IET Image Process., vol. 10, no. 6, pp. 448–455, 2016.

[2] A. Lensen, H. Al-Sahaf , M. Zhang, and B. Xue , "Genetic programming for region detection, feature extraction, feature construction and classification in image data," in Proc. Euro. Conf. Genetic Program., 2016, pp. 51–67.

[3] L. Yu, H. Chen , Q. Dou , J. Qin, and P.-A. Heng , "Automated melanoma recognition in dermoscopy images via very deep residual networks," IEEE Trans. Med. Image., vol. 36, no. 4, pp. 994–1004, 2017.

[4] Q.U.Ain, B.Xue,H.Al-sahaf ,and M.Zhang, "Genetic programming for skin cancer detection in dermoscopicimages,"inProc.Congr.Evol.Comput., 2017, pp. 2420–2427.

[5] M. Iqbal, B. Xue, H. Al-Sahaf, and M. Zhang, "Cross-domain reuse of extracted knowledge in genetic programming for image classification," IEEE Trans. Evol. Comput., vol. 21, no. 4, pp. 569–587, Aug. 2017.

[6] Q. Ul Ain, H. Al-Sahaf, B. Xue, and M. Zhang, "A multi-tree genetic programming representation for melanoma detection using local and global features," in Proc. 31st Australas. Joint Conf. Artif. Intell., 2018, pp. 111–123.