Ilkogretim Online - Elementary Education Online, 2020; Vol 19 (Issue 3): pp. 4634-4641
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2020.03.735617

Discernment Of Skin Cancer Using Machine Learning

Ms.S.Premalatha Assistant Professor , Dept of EIE, Sri Sairam Engineering College,Chennai-44,

ABSTRACT

The objective of the project is to classify skin lesion and cancer. A novel method is proposed that combines color and texture for the segmentation of skin lesions from unaffected skin region in an image. This project proposes a novel approach for classification of skin lesion and cancer images. The proposed work comprises of Pre-Processing, Segmentation, Feature extraction and Classification. In the Pre- Processing stage, Anisotropic diffusion Filter is implemented to remove noise and undesired structures from the images. In the Segmentation stage Fast Fuzzy C Means clustering method is implemented in order to acquire a contour by means of the gradient flow that minimizes an energy function with a distance regularization term and an external energy that drives the motion of the zero level set toward desired locations. The Gray level Co-occurrence Matrix (GLCM) and bandlet transform are used to estimate the features of the segmented image. The convolutional neural network classifier is employed for the classification task, utilizing feature vectors derived from gray level co- occurrence (GLCM) features. The classification results are evaluated with the use of accuracy, sensitivity and specificity. An automated Matlab tool is developed for classification of skin lesion and cancer.

Keywords

Dermoscopic Digital images, Feature extraction, gray level co -occurrence matrix and bandlet transform.

I INTRODUCTION

The majority of skin cancer deaths are due to malignant melanoma. It is considered as one of the most dangerous cancers. In its early stages, malignant melanoma is completely curable with a simple biopsy. An early detection is the best solution to improve skin cancer prognostic. Medical imaging such as dermoscopy and standard camera images is the most suitable tools available to diagnose melanoma at early stages. To help radiologists in the diagnosis of melanoma cases, there is a strong need to develop computer aided diagnosis systems. The accurate segmentation and classification of pigment skin lesions still remains a challenging task due to the various colors and structures developed randomly inside the lesions.

II PREPROCESSING

To analyze skin lesions, it is necessary to accurately locate and isolate the lesions. The description of the border aspect appears to be an important feature for clinical judgment. The variation in color signifies the malignancy of the lesion. Hence effective discrimination of the skin lesions was based on the distribution of texture and color features in this project. A pre-processing step was adopted to remove the influence of skin lesions in the surrounding regions and also the influence of small structures, hairs, bubbles etc.

III IMAGE RESTORATION

The simplest and best investigated diffusion method for smoothing images is to apply a linear diffusion process. We focus on the relation between linear diffusion filtering and the convolution with a Gaussian, analyze its smoothing properties for the image as well as its derivatives, and review the fundamental properties of the Gaussian scale-space induced by linear diffusion filtering.

IV IMAGE ENHANCEMENT

ADAPTIVE MEAN ADJUSTMENT (AMA)

AMA is a computer image processing technique used to improve contrast in images. It differs from ordinary histogram equalization in the respect that the adaptive method computes several histograms, each corresponding to a distinct section of the image, and uses them to redistribute the lightness values of the image. It is therefore suitable for improving the local contrast of an image and bringing out more detail.

However, Adaptive histogram equalization has a tendency to over- amplify noise in relatively homogeneous regions of an image. A variant of adaptive histogram equalization called AMA algorithm prevents this by limiting the amplification.

Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR)

These are used to measure the degree of image distortion because these can represent the overall gray-value error contained in the entire image. MSE measures gray-level difference between pixels of the ideal and the distorted images without considering correlation between neighboring pixels. PSNR is defined as (10*log10) of the ratio of the peak signal energy to the MSE observed between the original image (I) and the filtered image (I').

The formula for weighted MSE is,

$$\sum^{M} \quad \sum^{N} \quad |(i, \quad j) \quad - \quad I'(i, \quad j)|\wedge 2$$

Where,

i=1

j=1

IMAGE

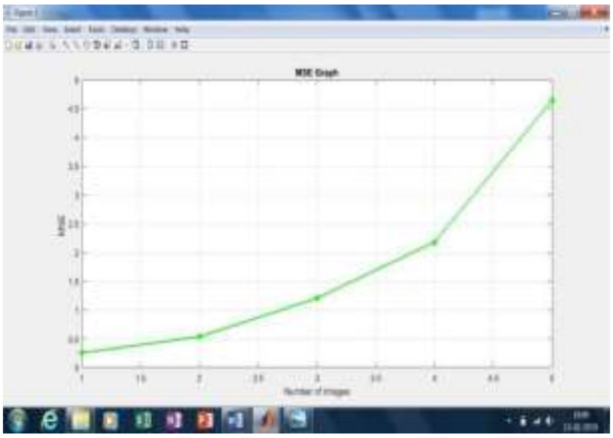
SEGMENTATION

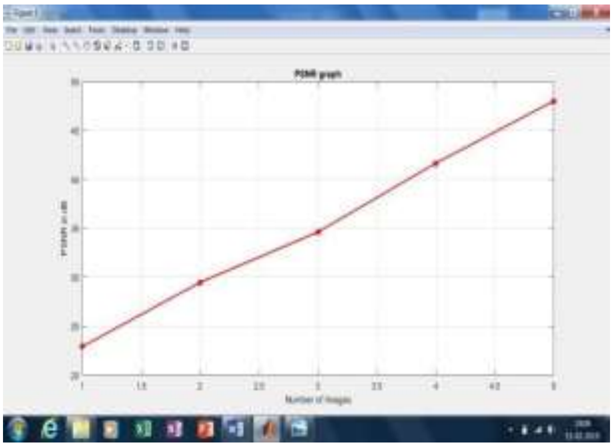
M, N – width and height of original image,

I(i,j) – original image, I'(i,j) – filtered image, The formula for PSNR is,

$$10 * \log10 (MAX^2 / MSE)$$

Where, the maximum fluctuation in the input image data type is divided by MSE. The following graphs shows the plot of MSE and PSNR respectively.





FAST FUZZY C MEANS (FFCM)

In this project FAST FUZZY C MEANS (FFCM) segmentation algorithm will be implemented by initializing standard membership values (computed by template averaging) to obtain standard segmentation result for each and every image. The number of clusters, initial standard intensity values and the feature vector were given as the input. After the completion of initial clustering process, the clusters were updated which in turn updates the cluster weights and membership degree. An objective function will determine that up to which the clusters should be updated.

IV FEATURE EXTRACTION

The Following GLCM features were extracted in our project work: Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster Shade, Dissimilarity Energy, Entropy, Homogeneity, Maximum probability, Sum of squares, Sum average, Sum variance, Sum entropy, Difference variance, Difference entropy, Information measure of correlation, normalized inverse difference. The 2D band-let transform is used to convert the spatial domain skin image into frequency domain image to estimate real and imaginary part of the image.

V DESCRIPTION

OF EXTRACTED FEATURES

(i) CONTRAST

Contrast returns a measure of the intensity between a pixel and its neighbor over the whole image. Contrast is 0 for a constant image. Contrast is calculated by the equation,

C = (Sumi(Sum j (i -j)^{2}Cij))

(ii) ENERGY

Energy returns the sum of squared elements in gray level co -occurrence Matrix (GLCM). Energy is also known as uniformity. The range of energy is [0,1]. Energy is 1 for constant image . The formula for finding energy in given equation

E = (Sumi{ Sumj(Cij^2)}

(iii) CORRELATION

iσj}

Correlation returns a measure of how correlated a pixel is to its neighbor over the whole image. The range of correlation is [-1,1].

Correlation is calculated by

Corr = Sum ij{(i-μ) *(j-μ) *P (i,j) /σ

(iv) HOMOGENEITY

Homogeneit**y** returns a value that measures the Closeness of elements in GLCM to the GLCM diagonal. The range of Homogeneity is [0,1]. Homogeneity is 1 for a diagonal GLCM. The homogeneity is evaluated using the equation

μ = {Sumi(Sumj[Cij/ (1+(abs(i-j)))]

(v) ENTROPY

Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image.

(vi) VARIANCE

Variance is used to find how each pixel varies from the neighboring pixel (or center pixel) and is used in classifying different regions.

(vii) CROSS CORRELATION

Cross-correlation is the comparison of twodifferent time series to detect if there isa correlation between metrics with the
same maximum and minimum value

VI CLASSIFICATION

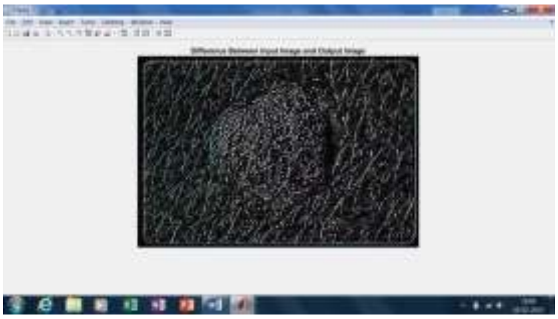
CONVOLUTION NEURAL NETWORK

Convolution Neural Network (CNN) is commonly known as neural network which is a mathematical motivated by the structure of biological neural networks. A neural network consists of an interconnected group of artificial neurons. This work describes the use of neural network in classification of medical images such as skin cancer and lesion , where the input units represents the feature vector and the output units represents the pattern class which has to be classify. Each input vector (feature vector) is given to the input layer, and output of each unit is corresponding element in the vector. Each hidden units calculates the weighted sum of its input to outline its scalar a net activation. The product of input vector and weight matrix at the hidden layer is generally called net activation function. Convolution Neural network is a useful tool for skin cancer classification. The advantage of parallel processing in deep learning and their ability to classify the data based on features provides a promising platform for medical image classification. Traditional sequential processing techniques have limitations for implementing pattern recognition problems in terms of flexibility and cost whereas back propagation neural networks perform the processing task by training the features calculated from the training signals instead of programming in a manner similar to the way human brain learns. The CNN works based on the trained features fed to the hidden layer of the neural network through the training sequence. The various features are calculated and fed to the hidden layer through the sigmoid function based on the weight vector generated by the input layer of the CNN.

VII RESULTS

(i) INPUT IMAGE

(ii) RESTORED IMAGE

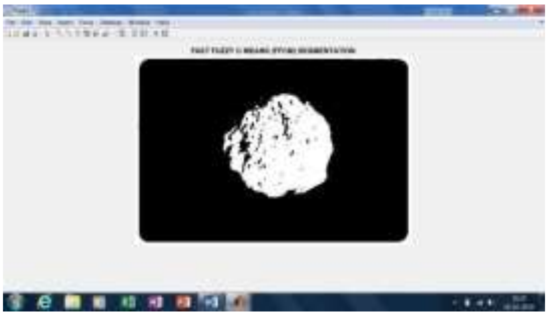


(iii) ENHANCED IMAGE

(i) CLUSTERED IMAGE

(V) THRESHOLDING

(vii) SEGMENTED IMAGE



(viii) EXTRACTED,FEATURES

FEATURES	VALUES
Contrast	1.2416

Correlation	0.9201
Energy	0.6583
Homogeneit y	0.9778
Sum entropy	0.5915
Sum variance	47.2404

VIII FUTURE SCOPE

This project can also be implemented for all type of cancers. This helps patients undergo painless diagnosis and reduces need of skin biopsy for classification of skin lesion and cancer.

IX REFERENCES

[1] N. Howlader, A. M. Noone, M. Krapcho, J. Garshell, N. Neyman, S.F. Altekruse, C. L. Kosary, M. Yu, J. Ruhl, Z. Tatalovich, H. Cho, A.Mariotto, D. R. Lewis, H. S. Chen, E. J. Feuer, and K. A. Cronin, "SEER cancer statistics review, 1975-2010," Nat. Cancer Inst., Bethesda, MD,USA, Tech. Rep., 2013

[2] A. F. Jerants, J. T. Johnson, C. D. Sheridan, and T. J. Caffrey, "Early detection and treatment of skin cancer," Amer. Family Phys., vol. 62, no. 2,pp. 1–6, Jul. 2000.

[3] Public Health Agency of Canada. (2013). Melanoma skin cancer. [Online]. Available:http://www.phac- aspc.gc.ca/cd-mc/cancer/melanoma skin cancer-cancer peaumelanome-eng.php

[4] A. Jemal, M. Saraiya, P. Patel, S. S. Cherala, J. Barnholtz-Sloan, J. Kim, C. L. Wiggins, and P. A. Wingo, "Recent trends in cutaneous melanoma incidence and death rates in the united states, 1992-2006," J. Amer. Acad. Dermatol., vol. 65, no. 5, pp. S17.e1– S17.e11, Nov. 2011.

[5] K. A. Freedberg, A. C. Geller, D. R. Miller, R. A. Lew, and H. K. Koh, "Screening for malignant melanoma: A cost-effectiveness analysis," J. Amer. Acad. Dermatol., vol. 41, no. 5, pt. 1, pp. 738–745, Nov. 1999.

[6] Rafael C Gonzalez ,Richard E Woods "Digital Image Processing" published by Prentice Hall. 2009.