

Analyzing Human Skin Texture using Machine Learning Approaches

M. Preethi
Research Scholar
PSGR Krishnammal College for Women
Coimbatore

K. Sathiyakumari
Assistant Professor
PSGR Krishnammal College for Women
Coimbatore

ABSTRACT

Analysis of skin texture is very useful for creation and development of cosmetic products, skin texture modeling, and face recognition in security applications and also computer-assisted diagnosis in dermatology. Several types of skin diseases are increasing human begins daily life; to deal with an effective and also very important manner the disease must be diagnosed properly. Skin texture analysis is one of the major problems in the field of medical diagnosis for finding skin diseases. Hence, the texture of skin is analyzed based on various features and characteristics so that the inconsistencies can be avoided during the treatment. The main goal of this study was to examine the texture of the human skin by image processing method. The skin properties like skin oiliness, dryness, pigmentation, fungus, infection, allergic symptoms and itching kind of problems association with skin texture profile is debated in the proposed work. Skin images are pre-processed using various pre-processing techniques and the Texture Filtering method is used for segment the skin textures so it can easy to identifying the skin properties accurately. Finally machine learning techniques are used to analyze and categorize the skin textures based on the texture and shape features. The experimental result shows that Decision Tree algorithm outperforms well in categorizing skin textures.

Keywords

Machinelearning, Skin Texture Analysis, Matlab2012 Software.

1. INTRODUCTION

The basis on which the human is able to distinguish between surfaces and objects is texture analysis. The skin layer consistencies have an intimate relationship with the individual's hormones, hydration, diet and any allergic symptoms of the body. Several properties show a relationship with the skin texture e.g. skin fungus, dryness, allergic symptoms etc.

Texture analysis is considered as a process of the spatial difference in pixel intensities (gray values), is effective in a range of functions moreover is being a condition of powerful research by a lot of research workers. Texture is the most essential image sign in recognizing the different kinds of standardized segments. This is really called texture analysis [4].

Large amount of work has been performed on analysis of the skin texture up to date. Previously the image is analyzed for extracting hemoglobin and melanin components by independent requirement reports. The representation have been synthesized for the comparison and the texture contents are also rotten for filtering the representation of skin layers into the initial structure along with the feature vector based on which the pixel change is observed to find out the pattern and structure [3]. The filtering technique is always invented which

is really used on dermoscopic skin representation in a non-linear approach and also permits exclusive image filtration

The attribute is seriously enviable considering in many instances of computer enabled diagnostic, get into the images need to be pre-process and this can results in unnecessary artifacts as well as may need natural confirmation [4].

2. RELATED WORK

Damanpreet Kaur, Prabhneet Sandhu. The skin texture images used to analysing diagnose of the skin diseases. The experimental result has been present effective separation of skin appearance by wavelet evaluation and also reasonable synthesized images.

Anil Kumar Mittra et al [6]. In this particular work to approach the automatic identification of Skin Diseases using GLCM features. In this experimental result is 96.6% obtained using a multilayer Perceptron (MLP) as a classifier.

Lei Huang et al [7]. Extended the fundamental MSER to each color and also structure evaluation to decrease the skin layer similar to areas. Additionally, the flexible effective radiance together with the chrominance, skin identification is going to be used to change the skin color model to each and every image. The procedure is made of obtained ensuring overall performance. This explains an exciting collection of datasets. The True positive rate is 81.2% under false positive rate is 8.2%.

Chen Guannan et al [8], evaluation algorithms of mean- shift are carried out on reports used to consistency specification. At the end of a composite formulation is provided for quantitative evaluation images. In this work, explain the texture features and also analytic given that paper perform well on specific the regular structure from scar mark structure including effectively using on their boundary's.

Parekh R [9]. Mentioned on automatic method for reorganization of Skin Disease problems by evaluates Skin consistency visuals applying consistency reliability methods depending on Gray level co-occurrence matrix as well as wavelet decomposition matrices.

Al-Mohair H.K et al [10]. An extensive equivalent analysis using the Multilayer Perceptron reproduction neural network (ML P), which may be universal classifier, is performed to understand the overall efficiency of various color-spaces for skin identification. The most productive color space applying color (or) color- texture qualities individually.

3. METHODOLOGY

3.1 K-Nearest Neighbor

KNN is a case-based learning method, which preserves all the training data for classification. Being a lazy learning method excludes it in many applications. One way to increase its

efficiency is to find some representatives of the whole training data for classification, viz. creating an inductive learning model from the training dataset and using this model for classification.

Algorithm

Input: The feature weight vector weight, the train data feature vector with numeric features, value of k and the test data.

Output: The predicted category of the test data

Method:

1. Find the k nearest neighbors to the test data using the feature weighted distance formula as

$$\text{distance}(X, Y) = \sqrt{\sum_{i=1}^n \text{weight}_i (x_i - y_i)^2}$$

2. A certain distance weighted voting on the type of the k nearest neighbors is used to predict the category of the test data.

a. Consider an excess weight to every one of the k nearest neighbor as given that below

$$w_i = \begin{cases} \frac{d(x, x_k^{NN}) - d(x, x_i^{NN})}{d(x, x_k^{NN}) - d(x, x_1^{NN})} & \text{if } d(x, x_k^{NN}) \neq d(x, x_i^{NN}) \\ 1 & \text{if } d(x, x_k^{NN}) = d(x, x_i^{NN}) \end{cases}$$

Where x' is the test data, $xkNN$ is the k th nearest neighbor to the test data, $x1NN$ is the first nearest neighbor to the test data and $xiNN$ is the i th

y' of the test data as

$$\hat{y} = \underset{y \in \{y_1^{NN}, \dots, y_i^{NN}\}}{\text{argmax}} w_i \times \delta(y = y_i^{NN})$$

The function

$$\delta(y = y_i^{NN}) = \begin{cases} 1, & y = y_i^{NN} \\ 0, & y \neq y_i^{NN} \end{cases}$$

Where y_i^{NN} is the class label of the i th nearest neighbor, y is the set of all class labels.

3.2 Naive Bayes

The Naive Bayes technique produces the common theory that most of the features are separated. This leads to a much simpler, though surprisingly effective classifier in practice. The independence assumption immediately implies that the likelihood may be decomposed into an item of dimension-wise possibilities:

$$P(X|c_i) = P(x_1, x_2, \dots, x_d | c_i) = \prod_{j=1}^d P(x_j | c_i)$$

Algorithm

Step1. Let there be ‘m’ Classes: C1, C2, C3... Cm

Step2. Naive Bayes classifier predicts X belongs to Program Ci iff

$$P(C_i/X) > P(C_j/X) \text{ for } 1 \leq j \leq m, j \neq i$$

Step3. Maximum Posteriori Hypothesis

$$P(C_i/X) = P(X/C_i) P(C_i) / P(X)$$

Step4. Naïve Assumption of “quality conditional Independence”

$$P(X / .Ci) = \prod_{k=1}^n P(x_k / Ci)$$

$$P(X/Ci) = P(x1/Ci) * P(x2/Ci) * \dots * P(xn/ Ci)$$

3.3 Decision Tree

Decision tree produces recognition or regression products through a tree construction. It splits a dataset into smaller subsets while at the same time connected decision tree is incrementally developed. The final result is a decision tree with decision nodes and leaf nodes.

Algorithm

The core algorithm for building decision trees called ID3 presented by J. R. Quinlan which works as top-down, greedy search through the space of possible branches with no backtracking. It works with Entropy including Information and facts obtain to develop a decision tree.

ID3 (D; Attributes; Target ())

Step 1: t = create Node

Step 2: $\forall (x, c(x)) \in D: c(x) = 1$ THEN label (t) = ‘+’,
Return (t) ENDI

Step 3: IF $\forall (x, c(x)) \in D: c(x) = 0$ THEN label (t) = ‘-’,
Return (t) END

Step 4: label (t) = most Common Class (D; Target

Step 5: IF Attributes = \emptyset ; THEN return (t) END

Step 6: $A^* = \text{argmax } A \in A$ attributes (information Gain (D, A)

Step 7: FOREACH a $\in A^*$ DO

$$D_a = \{(x, c(x)) \in D: x|_{A^*} = a\}$$

IF $D_a = \emptyset$ THEN
t’ = create Node ()
Label (t’) = m

MostCommonClass (D, Target)

Create Edge (t, a, t’)

ELSE

Create Edge (t, a, ID3 (Da, Attributes \ {A*}, Target))

ENDIF

ENDDO

Step 8: return (t) Where, X- is a set of feature vectors, also called feature space.

C is a set of classes.

C: X! C is the ideal classifier for X.

$D = f \{(x_1, c(x_1)) \dots (X_n, c(x_n))\} \subseteq X \times C$ is a set of examples.

4. EXPERIMENTL RESULTS

4.1 Data acquisition

The data collecting is an important task in any machine learning activity. The images are related to seven types of human skin texture have been collected from Freepik.com.

Freepik.com is a privately-held company which focuses on the development of a medical and all other fields of images. Seven sets of images namely skin oiliness, dryness, pigmentation, fungus, infection, allergic symptoms and itching kind of problems, each consisting of 50 images are taken into consideration.

4.2 Preprocessing

Skin images are collected using the UV camera in order to get the deep skin images. The acquired images are stored in jpeg forms and are read in mat lab using the command imread (). Then resized using imresize () function. After resize the image is converted to gray image using rgb2gray () function. Image enhancement is done by using histogram techniques. Followed by enhancement process the segmentation, feature extraction and classification is performed.

4.3 Segmentation

Image segmentation is the process that subdivides an image into its essential parts or objects. The scope to which this subdivision is accomplished will depend on the segmentation need to avoid while the items of awareness in an application are generally unavailable. Image segmentation algorithms are used for image segmentation. In this segmentation phase, Texture Segmentation Using Texture Filters. Entropyfill and rough mask functions are used in segmenting the skin texture image. The result of segmented image is shown in Figure 1.

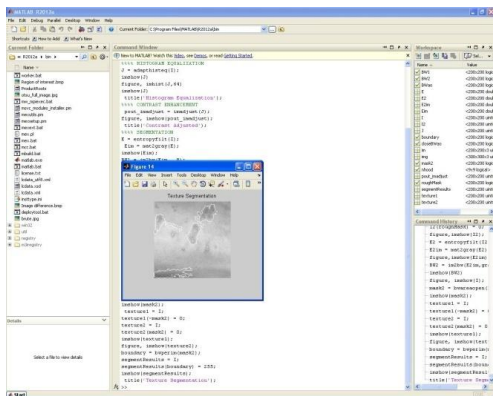


Fig: 1 Result of segmented image

4.4 Feature extraction

Two types of features such as shape and texture features are extracted here in order to extract the prominent features of an image. Shape features like Count, Sensitivity, Intensity Threshold, and Maximum Corner Count and Corner Threshold are extracted using SURF algorithm. SURF (Speeded up Robust Features) is a robust local feature detector used in this approach to extract relevant features and descriptors from images. Texture features are extracted using statsture function. This function looks at several texture features such as smoothness, third moment, or skewness, uniformity and entropy, or randomness.

4.5 Classification

The experiments have been carried out by implementing machine learning techniques using MATLAB 2012a platform. There are three different supervised classification algorithms are used namely Decision Tree (DT), Naive Bayes (NB), and K-Nearest Neighbor (KNN). The model is trained and tested using two different samples like training set and testing set. Finally 10-fold cross validation is performed to obtain better accuracy using crossval () function. The performance of the

classifier is evaluated in terms of accurate. The classification result is shown in the following Figure 2.

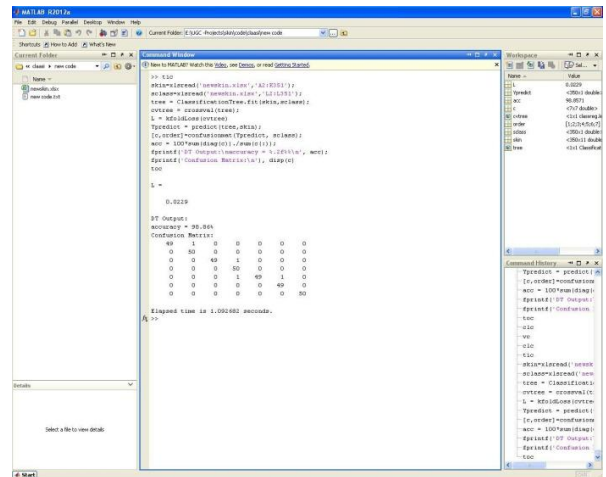


Fig: 2 Classification result

4.6 Result

The comparative results indicate that Decision Tree based classification model yields a better performance when compared to other models. The performance of the classifiers are assessed using measures such as accuracy, precision, recall, f-measure and time taken to build the model are shown in Table I.

Table 1 Predictive performance of classifiers

Classifier	DT	NB	KNN
Precision	0.9816	0.97340	0.8817
Recall	0.9886	0.9742	0.8829
F-Measure	0.9883	0.9742	0.8829
Accuracy	98.86%	97.42%	88.29%

Table I and Figure. 3 show that Decision Tree outperforms KNN and Naïve Bayes in terms of Precision, Recall, F-measures and Accuracy. 98.86% accuracy is achieved by Decision Tree whereas KNN and Naïve Bayes show only 88.29% and 97.42% respectively.

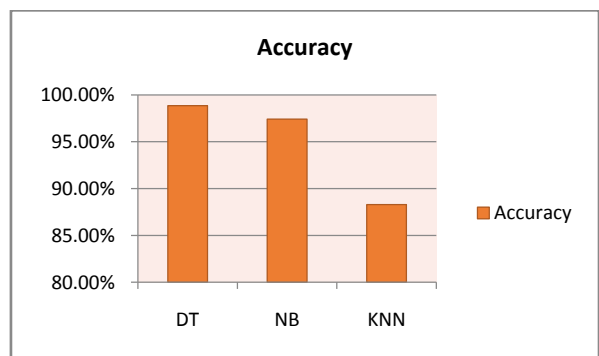


Fig: 3 Comparison of accuracy for various Classifiers

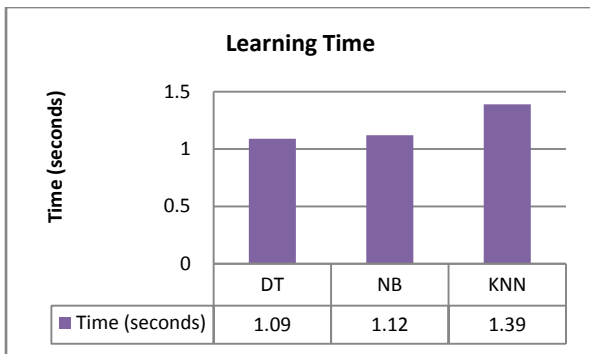


Fig: 4 Learning time for different classifiers

5. CONCLUSION

The main focus of this work is on analyzing the texture of skin thereby using it to diagnose the skin diseases, cosmetic products, skin texture modeling and security purposes. Various skin textures can be analyzed based on the combination of texture and shape feature vectors like contrast, correlation, energy, homogeneity, Count, Sensitivity, Intensity Threshold, MaximumCornerCount and CornerThreshold. The machine algorithms such as Decision Tree (DT), Naive Bayes (NB), and K-Nearest Neighbor (KNN) are used to classify the skin texture types.

From the experimental results discussed above, it infers that the Decision Tree classification can serve as an effective classifier in identifying skin textures accurately with above 98.86% of accuracy. The future work will be based on developing algorithms to identify various other skin textures for diagnoses skin disease possibilities, to improve the overall efficiency and also to supplementary reduce the computational time.

6. REFERENCES

- [1] Ranjanparekh, "Using Texture Analysis for Medical Diagnosis" Jadavpur University, India, 1070-986x/12, published by IEEE computer society, vol. 177, May 2011.
- [2] Kaur D., Sandhu P. Human skin Texture Analysis using image Processing Techniques, IJSR, and ISSN: 2319-7064, 2012.
- [3] Norimichi Tsumura, Nobutoshi Ojima, Kayoko Sato, "Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin", IEEE-2012
- [4] NeilT.Clancy A, Martin J. Leahya, Gert E. Nilssonb, Chris Andersonc, "Analysis of skin recovery from mechanical indentation using diffuse lighting and digital imaging", Proc. of SPIE-OSA Biomedical Optics, SPIE Vol. 6629, 66291G, © 2007 SPIE-OSA .1605-7422/07/\$18
- [5] YuantingGu and Enhua Wu, "Feature Analysis and Texture Synthesis", 978 -1-4244-1579-3/07/\$25.00 C) 2007 IEEE.
- [6] Anil Kumar Mitra, Dr.Ranjan Parekh, "Automated Detection of Skin Disease Using Texture", International Journal of Engineering Science and Technology (IJEST).
- [7] Lei Huang, Tian Xia; Yongdong Zhang, Shouxun Lin, "Human skin detection in images by MSER analysis", IEEE International Conference on Image Processing (ICIP), 2011, pp.1257 – 1260.
- [8] Chen Guannan, Xie Zhiming , Lin Juqiang ,Chen Rong ,Yang Kuntao , "Texture analysis on two-photon excited microscopic images of human skin hypertrophic scar tissue", PhotonicsGlobal@Singapore, 2008, IPGC 2008. IEEE, pp.1 – 4
- [9] Al-Mohair H.K, Mohamad-Saleh, J, Suandi, S.A, "Color space selection for human skin detection using color-texture features and neural networks", International Conference on Computer and Information Sciences (ICCOINS), IEEE 2014, pp. 1 - 6
- [10] J.R. Quinlan, "C4.5, Programs for Machine Learning", Morgan Kaufmann, San Mateo, Ca, 1993.
- [11] Chun-Nan Hsu, Hung-Ju Huang, Tsu-Tsung Wong, "Why Discretization works for Naïve Bayesian Classifiers", 17th ICML, pp 309-406, 2000.
- [12] P. Domingos, M. Pazzani, "Beyond independence: conditions for the optimality of the simple Bayesian Classifier", Machine Learning Proceedings of the Thirteenth International Conference, Morgan Kaufman, July 1996.
- [13] F. Anguilli, Fast Condensed Nearest Neighbor Rule, Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany, 2005.
- [14] Bhatia, N., Ashev, V. "Survey of Nearest Neighbor Techniques", International Journal of Computer Science and Information Security, Vol. 8, No 2, pp.1- 4, 2010.