Ilkogretim Online - Elementary Education Online, 2020; Vol 19 (Issue 2): pp. 2076-2085
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2020.02.696792

Melanoma Skin Cancer Detection Using Knn And Svm Classifier

Dr. K. Renganathan Professor and Head, Department of Electronics and Instrumentation Engineering, Sri Sairam Engineering College, Chennai, Tamilnadu, India

Dr. Vidyacharan Bhaskar Adjunct Faculty Instructor,Department of Electrical and Computer Engineering, San Francisco State University,San Francisco, CA 94132

J. Vishnuvardhan Director, Crion Technologies, Chennai, Tamilnadu, India

Abstract: Skin cancer is life threatening disease which causes human death. Abnormal growth of melanocytic cells causes a skin cancer. Due to malignancy feature skin cancer is also known as melanoma. Melanoma appears on the skin due to exposure of ultraviolet radiation and genetic factors. So melanoma lesion appears as black or brown in colour. Early detection of melanoma can cure completely. Biopsy is a traditional method for detecting skin cancer. This method is painful & invasive and requires laboratory testing so it is time consuming. Therefore, in order to solve the above stated issues computer aided diagnosis for skin cancer is needed. Computer aided diagnosis uses Dermoscopy for capturing the skin image. In this paper first pre-processing of the skin image is done. After pre-processing lesion part is segmented by using image segmentation technique which is followed by feature extraction in which unique features are extracted from segmented lesion. After feature extraction, classification by using support vector machine and K-nearest neighbour classifiers is performed for classifying the skin image as normal skin and melanoma skin cancer. The proposed system results shows that by comparing the result of support vector machine and K-nearest neighbour will gives optimum accuracy.

Keywords: Melenoma Skin Cancer, Segmentation, Support vector Machine, K-nearest neighbor

1. INTRODUCTION

The skin is the largest organ in the human body and has numerous potential abnormalities; there are about 1500 distinct skin diseases. We are relatively ignorant about the symptoms of the majority of these diseases although knowledge is rapidly increasing, however, that makes it a challenge for dermatologist to diagnose them.

Nowadays technologies have changed our day-to-day life in all aspects and the medical field is not an exception, many medical systems have been developed to help both

patients and doctors in different ways, starting from registration process ending with the use of technologies for diagnosing diseases.

Skin diseases rate has been increasing for past few decades, many of these diseases are very dangerous, particularly if not treated at early stages. In addition, dermatologist use variety of visual clues such as color, scaling and arrangement of lesions, the body site distribution and others, when these individual components are analyzed separately, the recognition of the disease can be quite complex that requiring high level of experience. Diagnosis by humans depends upon subjective judgment of the dermatologists so it's hardly reproducible, unlike computer aided diagnostic systems which are more realistic and reliable.

2. PROPOSED METHODOLOGY

In this paper the proposed diagnosis system mainly consists of following components:

- IMAGE ACQUISITION
- IMAGE PRE-PROCESSIGN
- IMAGE SEGMENTATION
- FEATURE EXTRACTION
- DATA MINING UNIT
- SUPPORT VECTOR MACHINE
- K-NEAREST NEIGHBOUR

The process flow of the above components for diagnosis is shown in Fig 1.

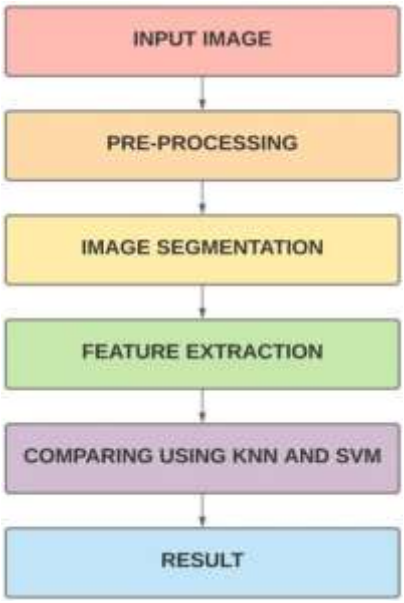


Fig 1.1 Process Flow chart

The support vector machine is a supervised learning model used for optimization. It is a unified framework in which different learning machine architecture can be generated through an appropriate choice of kernels. The principal used in SVM is statistical and structural risk minimization. The SVM is already a ready-to-use available classifier in MATLAB. After the feature extraction process, the extracted features are directly fed into the SVM classifier. The process involves two phases:

Training Phase: Huge images of eczema, impetigo, melanoma, and others are used for training.

Testing Phase: In this phase, test images are given to the classifier and the classifier uses knowledge gained during the training phase to classify the test image.

Fig 2: SVM Process

A K- NEAREST NEIGHBOUR (KNN) is slightly in variance with the multilayer perceptron. A KNN can have a single convolution layer or it can contain multiple convolution layers. These layers can be interconnected or pooled together. A convolution operation is performed on the input and then the results are passed to the further layers. Thus, due to this, the network can be deep but will contain only a few parameters. Due to this property, a Convolutional neural network shows effective results in image and video recognition, natural language processing, and recommender systems.

In SVM classifier, an image needs to be processed using image processing unit and then given for the classification to SVM classifier. KNN classifier is implemented in such a way where there is no need of image processing module. KNN classifier is a layered architecture where multiple layers perform various operations to train and test the image data. In this proposed solution, huge images are given to KNN classifier for training where images for training are given to Convolution2dLayer.

This is the first layer to extract the features from the input image. This layer applies a convolution operation and gives the result to the next layer and applies Convolutional filters to the input. It computes the dot product of the input and weights and then adds a

bias term. Then ReluLayer is introduced which is Rectified linear Unit Layer for handling nonlinearity in the network.

MaxPooling Layer reduces the dimensionality of image and is used to divide the input into rectangular regions and computes the maxima of each region. After this operation, Fully Connected Layer multiplies an input with weight matrix, adds bias vector and it is responsible for creating a model for classification layer by applying Softmax Layer. Softmax Layer is a logistic activation function which is used for multiclass classification. Finally Classification Layer will detect the affected area of image and gives the output.

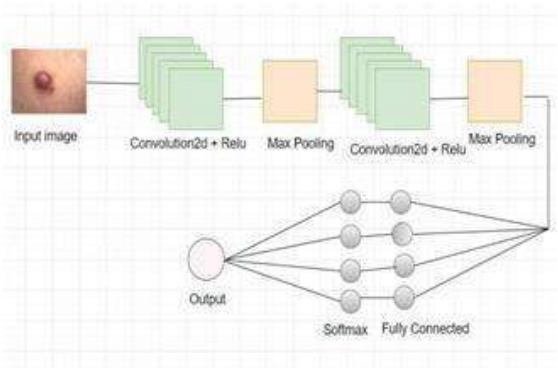


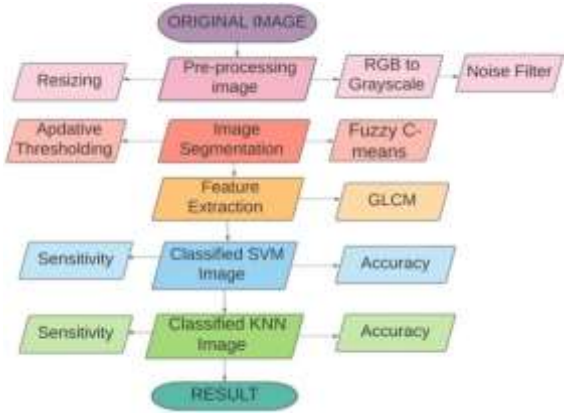
Fig 3: KNN Architecture

3. SOFTWARE REALIZATION

The proposed system has been designed to detect 3 types of skin diseases namely, Eczema, Melanoma and Acne. These classes were chosen because Acne and Eczema are of the most common skin diseases that have multiple effects on the patients on different aspects, they are painful in addition they have psychological effect caused by the changes that happen in skin specially for teenagers. Melanoma is the most dangerous form of skin cancer, most often caused by ultraviolet radiation from sunshine or tanning beds, If melanoma is recognized and treated early, it is almost always curable, but if it is not, the cancer can advance and spread to other parts of the body, where it becomes hard to treat and can be fatal.

For the training the system, 200 images for each of Acne, Eczema and Melanoma were collected and processed. The sample image collected in shown in Figure 4.

Fig 4: Image Sample



Images were converted from RGB – red, green and blue- type to gray scale images. Sample is shown in Figure 5.

Fig 5: Greyscale Image sample

Since Negative images are useful for enhancing white or grey detail embedded in dark regions of an image, then the gray images are converted to their negative, then power of two is applied to the negative image to darken the image. Before using Bag-of-Features as basis for training the data, list of techniques for classification was used, Table 1 shows the results of these models using jkkjk200 images as input for training and

evaluation, only Cross-Validation is applied.

Table 1: Different Classification Techniques

Learning Model	Associated accuracy
KNN	66%
Bag-of-Features	83%
HOG features with SVM classifier	55%

Bag of Features approach in computer vision in the past few decades has been used a lot in many applications. Bag of Features (BoF) methods have been applied to image classification, object detection, image retrieval, and even visual localization for robots.

In the proposed system, BoF approach is implemented to train our data, since it is used to classify images based on its texture. BoF approaches are characterized by the use of an order less collection of image features. Lacking any structure or spatial information this eliminates the effect of the water mark in our images which increase the accuracy compared with other learning models such as Convolutional Neural Networks and manually HOG features extraction along with SVM classification.

The algorithms flowchart for the process is shown in Figure 6.

Fig6: Algorithm Flowchart

4. RESULTS AND DISCUSSION

It is observed from below table that KNN Algorithm has near perfect accuracy and run time in detecting skin diseases shown in Table 2. The confusion matrix shows the percentage of error and accuracy and run time in classification. It also shows corrected and uncorrected results, true positives, false negatives and number of classes.

Table 2: Accuracy results

Disease	SVM Accuracy	KNN Accuracy	Runtime SVM (mS)	Runtime KNN (mS)
Eczem	98%	92%	421.19	19.24

a			8	
Impetigo	95%	91%	77.34	9.59
Melanoma	99%	93.4%	91.707	10.45
No Disease	92%	88.8%	114.15	21.76
Overall	96%	91.30%	176.0988	15.26

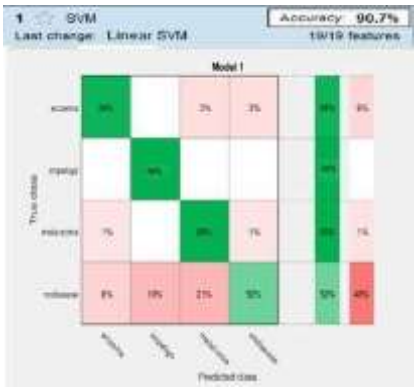


Fig 7: SVM Confusion Matrix

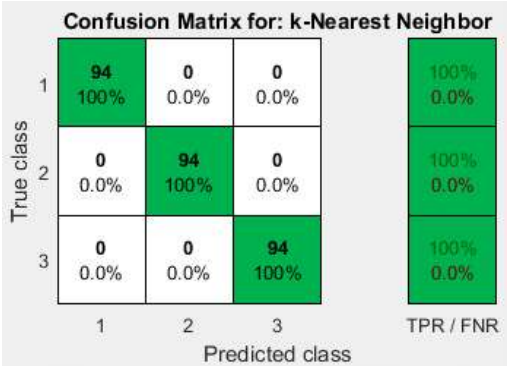


Fig 8: KNN Confusion Matrix

The output images obtained from the MATLAB simulation results are shown in Figures 9, 10 and 11 respectively.

Fig 9: Existing system

Fig 10: Proposed system

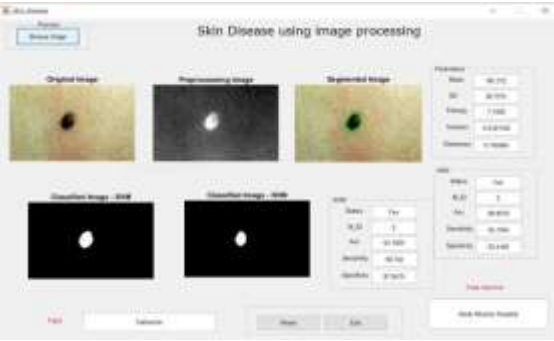


Fig 11: Output system

5. CONCLUSION AND FUTURE SCOPE

Difficulties in the diagnosing skin diseases arise because of the spreading of the skin diseases all over the world, which make it a challenge to the dermatologist to recognizethe different skin diseases easily, a computer aided system is proposed to resolve these difficulties, so a machine learning model based on bag of features algorithm is designed which use KNN as a classifier and SURF for feature extraction with an interface on mobile applications is developed for android devices, the core model is developed using MATLAB and the interface is developed using android studio on windows OS.

The developed system performs the required work with accuracy 94% within the dataset and 85% with the external data, and the integrated system is working properly, although the system works fine there are some points that bound the performance of the system such as that, there is a single processor of the system that is used to classify the images, which will reduce the performance of the whole system with the increase of the requests from multiple users.The proposed modifications of the skin diseases diagnosis system are generally to increase the performance of the system, resolve the system limitations, or to increase its capability.

So, there are several suggested modifications to both the system core model and the system mobile interface:

- Increase the training data used for training the model, not only in term of quantity but also obtaining more data from different resources namely collecting data from hospitals and healthcare centers, to increase the learning model generalization.
- Apply better preprocessing techniques to resolve the images distortions.
- Apply training data of more classes, that the model will be capable to recognize and diagnose more diseases.
- Develop a cross-platform application to work on different mobile platforms, which will increase the number of system users.
- Enhance the application interface to be more user friendly for better user experience policies.
- Develop a distributed system for skin diseases diagnosis to resolve the single server limitation, and increase the processing capabilities.
- Enhance the functionality of the system to be more useful by givingadvices for the users about the disease treatment
- The system may be combined with other medical systems to propose an integrated medical care services, the suggested system is TBibak.

REFERENCES

[1] C. Barata, M. E. Celebi, and J. S. Marques, “A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer,” IEEE J. Biomed. Heal. Informatics, vol. 23, no. 3, pp. 1096–1109, 2019.

[2] Victor and M. R. Ghalib, “Automatic detection and classification of skin cancer,” Int. J. Intell. Eng. Syst., vol. 10, no. 3, pp. 444–451, 2017.

[3] H. Iyatomi et al., “An improved Internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm,” Comput. Med. Imaging Graph., vol. 32, no. 7, pp. 566–579, 2008.

[4] Jaleel, S. Salim, and R. B. Aswin, “Computer aided detection of skin cancer,” Proc.

IEEE Int. Conf. Circuit, Power Comput. Technol. ICCPCT 2013, pp. 1137–1142, 2013.

[5] R. B. Oliveira, N. Marranghello, A. S. Pereira, and J. M. R. S. Tavares, "A computational approach for detecting pigmented skin lesions in macroscopic images," Expert Syst. Appl., vol. 61, pp. 53– 63, 2016.

[6] S. Afifi, H. GholamHosseini, and R. Sinha, "A system on chip for melanoma detection using FPGA-based SVM classifier," Microprocess. Microsyst., vol. 65, no. February 2019, pp. 57–68, 2019.

[7] Murugan, S. A. H. Nair, and K. P. S. Kumar, "Detection of Skin Cancer Using SVM, Random Forest and kNN Classifiers," J. Med. Syst., vol. 43, no. 8, 2019.

[8] P. Y. Hsiao, S. S. Chou, and F. C. Huang, "Generic 2-D gaussian smoothing filter for noisy image processing," IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON, vol. 2, no. 2, 2007.

[9] R. J. Hemalatha, T. R. Thamizhvani, A. J. A. Dhivya, J. E. Joseph, B. Babu, and R. Chandrasekaran, "Active Contour Based Segmentation Techniques for Medical Image Analysis," Med. Biol. Image Anal., 2018.

[10] S. S. Yasiran, S. Salleh, and R. Mahmud, "Haralick texture and invariant moments features for breast cancer classification," AIP Conf. Proc., vol. 1750, no. June, 2016.

[11] M. Martfnez-Arroyo and L. E. Sucar, "Learning an optimal naive Bayes classifier," Proc. - Int. Conf. Pattern Recognit., vol. 3, no. January 2006, pp. 1236–1239, 2006.

[12] D. Dumitru, "Prediction of recurrent events in breast cancer using the Naive Bayesian classification," Ann. Univ. Craiova, Math. Comp. Sci. Ser., vol. 36, no. 2, pp. 92–96, 2009.