Int. J. Nonlinear Anal. Appl. 13 (Special issue for ICDACT-2021), 61–69
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.6331

Fingernail analysis for early detection and diagnosis of diseases using machine learning techniques

K. Dhana Shree[a,*], P. Jayabal[b], A. Saran Kumar[c], S. Logeswari[c], K. Ranjeetha Priya[d]

[a]*Department of CSE, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India*

[b]*Department of Mathematics, Rathinam Technical Campus, Coimbatore, Tamil Nadu, India*

[c]*Department of CSE, Bannari Amman Institute of Technology, Erode, Tamil Nadu, India*

[d]*Department of Information Technology,Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu, India*

(Communicated by Mayank Sharma)

Abstract

Each and every human have unique fingernails. In the early days, the psychological conditions of the human body were reflected with the help of the growth situation of the surface of nails. It is possible to diagnose human nails and predict the disease. Predicting the disease at the early stage helps in preventing the disease. In this proposed work, the image of the nail is taken from a microscopic image. The lunula and nail plate are segmented effectively using the image pre-processing techniques. Histograms of oriented gradients and local binary patterns are used to capture the characteristic value. Once after pre-processing various features of the nails are extracted using various machine learning algorithms such as Support Vector Machines, Multiclass Support Vector Machine, Convolution Neural Network along with an Optimization algorithm named Ant Colony Optimization to improve the efficiency of classification.

Keywords: Local Binary Pattern (LBP), Block Chain Technology (BCT), Machine Learning (ML)
2020 MSC: 60J20

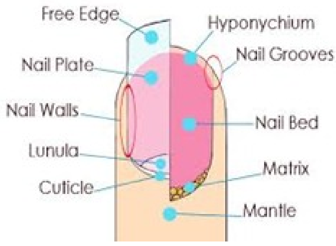
1 Introduction

Digitalization and usage of modern automated machines have made humans very lazy. This results in lack physical exercises which ultimately lead to poor health conditions. Now-a- days in this modern era, even children at very young ages are prompt to various diseases. Early detection and diagnosis of disease should be done to prevent the wide spread. Among all disease, it is highly complex to diagnose the dermatological diseases since they have shared characteristics and subjectivity of human interpretation. Human nail plays a vital role in various diseases identification and prediction. It is possible to predict many of the human diseases by just seeing the colour of the nail. A human having his nail pink in colour indicates that he is hale and healthy. Though various disease can be diagnosed using the colour of finger nails, the accuracy rate sometimes fails. This is mainly due to the colour assumptions made by humans through naked eye. Human eye has limitation in resolution and small amount of colour change in few pixels on nail would not be highlighted to human eyes which may lead to wrong result whereas it is possible for a machine

*Corresponding author
Email addresses: dhanu1304@gmail.com (K. Dhana Shree), kongu2008@gmail.com (P. Jayabal), sarankumar@bitsathy.ac.in (A. Saran Kumar), logeswari@bitsathy.ac.in (S. Logeswari), ranjeethapriya.k@gmail.com (K. Ranjeetha Priya)

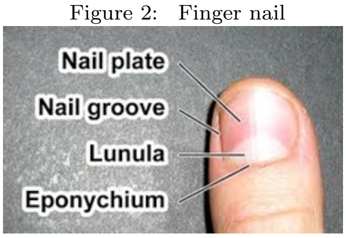
Received: August 2021 *Accepted:* January 2022

Figure 1: Structure of a finger nail



to recognize small colour changes on nail [3]. The health condition can be diagnosed using the nail's thickness, length of nails, colour and texture.

Figure 2: Finger nail



The structure of human nail is as follows:

- The nail matrix is the place where nail grows.
- The nail plate is the noticeable part of the nail.
- The nail assembles at the top of the nail bed.
- Lunula is the crescent-moon contour seen at the plinth of the nail plate.
- Nail folds are the meager skin furrows which hold the nail plate.
- The cuticle is the flutter of thin hankie over the plinth of the nail plate.

2 Related works

Vipra Sharma et al. [1] made a survey by taking the images of the nail and by extracting the features values and compared it with the feature values of a healthy nail. Initially the image was segmented and the feature values are obtained from the segmented area.

Trupti S Indi et al. [7] also made a survey to diagnose disease at early stage. He carried out the analyse using the WEKA tool. Few nail images were taken and trained. The nails images are tested and comparison is made with trained image. The results showed that there was exact match between the values of trained and tested images.

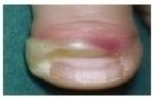
HardikPandit et al. [3] used the nail color for the analysis and predicted the disease. The human palm nail is analysed for predicting the probable disease. It was possible for HardikPandit et al to get accurate results and the limitations of human eye such as subjectivity and resolution power have been overcomed.

SukdeepKaur et al. [4] analysed various image segmentation techniques such as threshold segmentation, point transformation, watershed transformation and wavelet transform to get a clear segmented image. The logarthmic operator method in point transformation compressed the image by replacing each pixel value with its logarithm. Watershed transformation was used to separate two merged objects. To analyse horizontal, vertical and diagonal views of two dimensional image SukdeepKaur et al used watershed transformation method.

Nida M Zaitoun et al. [15] made study on block-based image segmentation techniques such as region-based and edge or boundary-based methods and made comparison between both the methods.

Muhammad [14] analysed various image processing techniques and concluded that there is no standard technique for segmenting an image. The main reason is that image segmentation process rely on various factors such as colour, pixel, texture, intensity, image content and problem domain. Muhammad also suggested that hybris solutions work

Table 1: Summary of Nail disease

Appearance of Nail	Nail Image	Medical name of the disease	Symptoms	Diagnosis
Redness around the nails	*Figure 3: Paronychia* 	Paronychia	Swelling, Itching, redness around the nails, pain	Diabetes
Terry's nail	*Figure 4: Terry's nail – Paronychia* 	Paronychia	Wrinkles around nails, aged look, shining nails	Diabetes, Organ failure, Liver cirrhosis, Hyperthyroidism, Renal failure, Liver cirrhosis, Congestive heart failure, Malnutrition, Thyroid problem
Thick and Overgrown Nails	*Figure 5: Onychogryphosis* 	Onychogryphosis	Nails grow fastly than normal pace	Arthritis or psoriasis
Black vertical stripe throughout the nail	*Figure 6: Melanonychia* 	Melanonychia	dark stripes that emerge along their nails	Type of skin cancer
Spoon nails	*Figure 7: Koilonychia* 	Koilonychia	Soft nails and spoon shaped	Hemochromatosis, (liver problem), heart disease and anemia
Brittle nails	*Figure 8: Onychoschizia* 	Onychoschizia	crumble or to become thin, frail and brittle	lack of vitamins nail psoriasis, thyroid problems or fungal nail infections

better for image segmentation process.

Nityash Bajpai et al. [1] proposed a new system for disaese detection, commomly known as Disease Detection System (DDS). The user's can take the image of human palm and nail and predict various disease in human nail. Using the nail's colour and texture observations are made and the disease in predicted.

3 Proposed work

The proposed work for disease prediction involves the following steps:

- Capturing the input image
- Image Segmentation
- Feature Extraction
- Classification
- Nail Disease Prediction

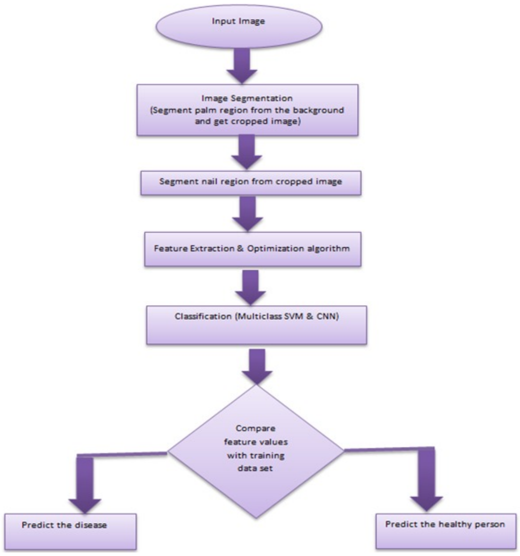


Figure 9: Flow diagram of the proposed work

Input image

The input image usually considered is the finger nail. The images will be captured through a high resolution digital camera. Sometimes, the input image can also be captured through a scanner and used.

Figure 10: Input Image

Image segmentation:

Once the input image is obtained, image segmentation has to be done.The palm can be segmented using the in built crop function. Image Segmentation is a process of segmenting or dividing the image into different parts until the region of interest is identified. In this proposed work, to segment the images we use region based segmentation method. It segmentes the image into number of regions or classes.

Feature extraction

The count of partitions in an image depends on number of features to be extracted.The process of redefining a large set of redundant data into a set of features of reduced dimension is known as Feature Extraction. Transforming the input data into the set of features is called feature extraction. In order to obtain the desired outcome few actions need to be performed. Once these actions are performed, the color, texture and shape-based features are extracted from whole images. Enhancing the edges uses Sobel operator to highlight the borders of the membranes and the nails . Canny edge detection can be used to obtain outputs with continuous edges. To keep the biased information, which the high-dimensional sample spaces are provided, good dimension reduction techniques are required. The main reason to go for dimensionality reduction technique is it improves accuracy and saves memory and time. Support vector machine technique is used to extract feature values.

Support Vector Machine works as follows:
Input: Training data set $X = \{X_1, X_2, ..X_m\}$ with class labels $Y \epsilon \{1..K\}$ and Testing data sets X_{test} , where K is the number of classes in the dataset
Output: Class labels for test data Y_{test}

- Select the regularization, kernel parameters for the classifier
- Construct SVM models. The model is trained with both positive and negative labels. Finally hyperplane is generated.
- Test sample xtest is assigned the class label that has the largest value of the decision function

Hyperplanes are drawn in order to perform classification in SVM. In two class classification, the hyperplane is equidistant from both the classes. Support vectors are the data instances that are used to define the hyperplane. The distance between hyperplane and the nearest support vector is defined as margin in SVM. Since large distance gives less error, the margin's distance should be as large as possible.In case if the margin is close, then sensitivity to noise is more. The equation to define the hyperplane is drawn and the margin are set to 0and ± 1 respectively. Here,the margins are defined as weight vector and bias.The input features given to SVM need to be reduced for getting appropriate results. This ultimately improves the efficiency of the results given by the algorithm.

Classification

The feature values are extracted and classification is done based upon the values. In order to enhance the features values or get optimized result, ant colony optimization technique is followed. The ant colony optimization metaheuristic is:

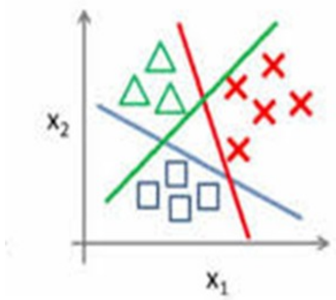
```
Input: ProblemSize , m, β, ρ, σ, q0
Output: Pbest
Pbest ← CreateHeuristicSolution(ProblemSize);
Pbest_cost ← Cost(Pbest);
Pheromone_init ← 1.0/ (ProblemSize * Pbestcost);
Pheromone ←InitializePheromone (Pheromone_init);
while StopCondition() do
    for i=1 to m do
        S_i ←ConstructSolution(Pheromone, ProblemSize, β, q0);
        S_iCost ← Cost(S_i );
        if S_iCost ≤ Pbest_cost then
            Pbest_cost ← S_iCost ;
            Pbest ← S_i ;
        end
    end
end
return Pbest;
```

The classification methods such as multiclass SVM and neural networks are considered and comparison is made between both the methods.

Multiclass SVM

Multiclass SVM is an efficient method for classifying the data. A classification basically involves two segments. Training and testing data along with data instances are the two segments. Each instance in the training set contains a target value and many other attributes. The target of SVM is to yield a mould that will forecast the data instances's target value in the testing set.

Figure 11: Multiclass SVM



SVM works well when dealing with simple features. Sometimes we need to extract complex features. The algorithm that bests suits for complex feature extraction is Convolution Neural Network.

Convolution neural network

The initial step in the Convolution Neural Network is performing convolution operation. The prime purpose of this operation is to extract the required features from the input images. Convolution operation is done between the input image and a filter of a specific size MxM. The next step involves using of pooling layer. This layer helps in reducing the size of convolved feature map so that computational costs can be reduced. Rectified Linear Unit Layer is used for performing linear functions. Once after performing linear functions, reducing the dimensions of feature map is mandatory to improve accuracy. This is done with the help of pooling layer. The data's need to be converted into 1-D array. This is done using flattening process. The fully connected layer will have inputs and that input is fed processed and becomes final output of the convolution network.

Figure 12: Convolution neural network to classify the disease

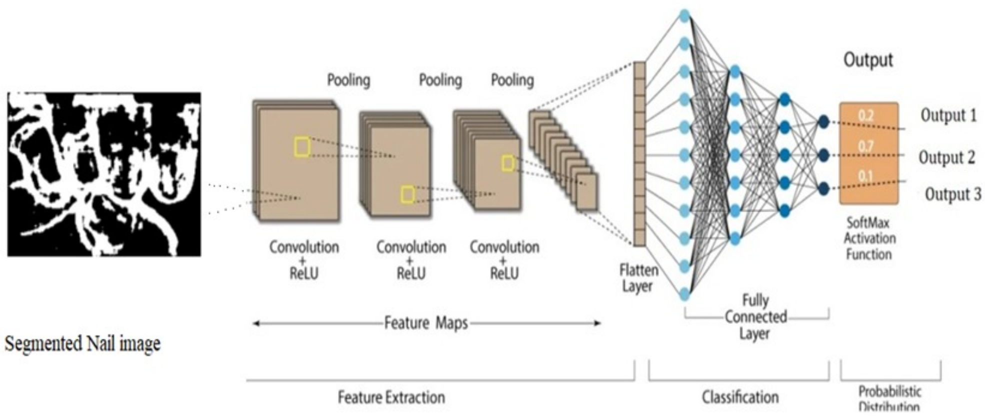


Figure 13: Pseudocode for convolution neural network

```
Input: Population size N
Output: Initial population P0
P0 → ϕ
while | P0| ≤ N do
        L=rand(0,1)
        lst → create linked list with L nodes
        for all nod € lst do
                r=rand(0,1)
                if r ≤ Ω0 then
                  nod.Type= 'Skip_Layer'
                  nod.F1=rand (0,1)
                  nod.F2=rand (0,1)
                else
                  nod.Type= 'Pooling_Layer'
                  tmp= rand(0,1)
                  if tmp ≤ 0.5 then
                    nod.P1=max
                  else
                    nod.P2=mean
                  end if
                end if
        end for
end while
P0 → P0 U lst
return P0
```

4 Experimental results and discussion

The system was developed with Python language and Tensor Flow backend. Proposed system used an input image size of 512x512, 0.0001as learning rate and 0.3 as weight decay rate to avoid overfitting. The optimizer used is ant colony optimizer as it is widely used for multi-class classification. The architecture was trained for 32 epochs. The architecture was trained and tested using two datasets from Kaggle and Messidor.
Publicly available Kaggle dataset consists of 34,589 nail images taken under varying conditions. Among them, 24,210 images were randomly chosen and divided into 70:15:15 ratio - train, validation, and test dataset. The 12,823 training images were trained and 5,188 were validated by the model. The 850 samples from the Messidor were split in the same pattern of 70:15:15. The 595 images for train data, 127 images for validation data, and 127 images for test data were taken. Two algorithms for classification namely multiclass SVM and Convolution neural network were used. The two algorithms are compared based on the following parameters:

- Accuracy
- Precision
- Sensitivity
- F1 Score

The metrics used for system evaluation are

where TP, TN, FP, FN are True Positive (TP) is classes with prediction true and ground truth also true. False Positive (FP) is classes with prediction true but ground truth is false. True Negative (TN) is classes with prediction false and ground truth also no. False Negative (FN) is classes with prediction false but ground truth is true.

The table below shows the calculated values based on the different metrics for Multiclass SVM and Convolution Neural Network:

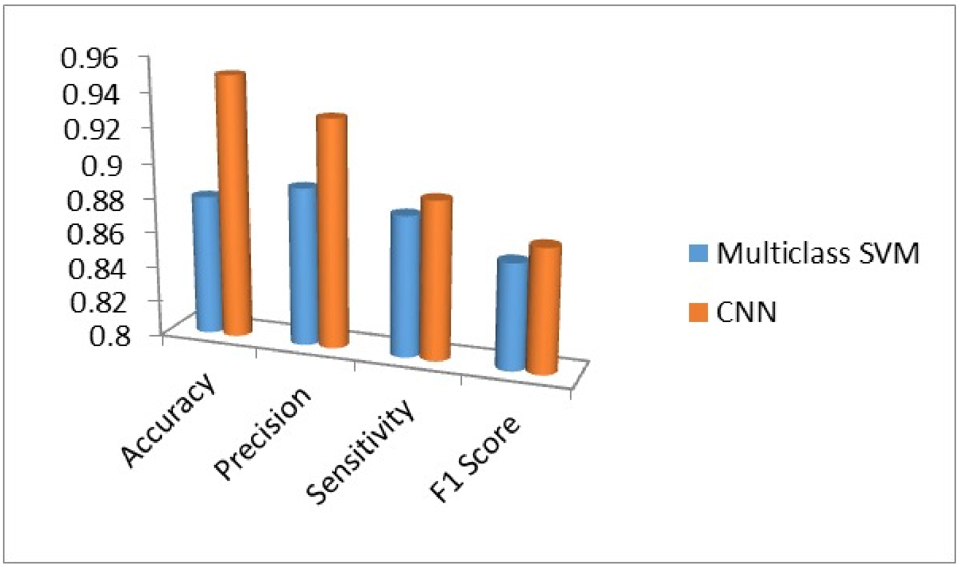
Metrics	Proposed Models	
	Multiclass SVM	Convolution Neural Network
Accuracy	0.88	0.95
Precision	0.89	0.93
Sensitivity	0.88	0.89
F1 Score	0.86	0.87

Metrics	Proposed Models	
	Multiclass SVM	Convolution Neural Network
Accuracy	0.88	0.95
Precision	0.89	0.93
Sensitivity	0.88	0.89
F1 Score	0.86	0.87

Table 2: Metrics values for multiclass SVM and CNN

The comparison of the two algorithms for the above parameters is plotted:

Figure 14: Comparison of Multiclass SVM and CNN



From the values calculated using different metrics and plot obtained, it is concluded that Convolution Neural Network works well for classifying disease from nail images and overall accuracy is 94.673% .

5 Conclusion and future work

In the above proposed work, the feature extraction of a nail image followed by classification was carried out. Security and privacy is still challenging when working with patient's data. Hence future work will focus on using a highly anonymous block chain technology to ensure privacy of data and manage the data used by user's through the block chain, this keeps record of any change or manipulation in the dataset and therefore the data can be kept highly secured.

Acknowledgment

We would like to extend our sincere thanks to our supervisor Dr. S.Logeswari for constantly motivating us to do the work and also extend our whole hearted thanks to Principal and Head of Department of CSE, Sri Ramakrishna Engineering College for the constant support in completing this work.

References

[1] N. Bajpai, R. Alawadhi, A. Thakare, S. Avhad and S. Gandhat, *Automated prediction system for various health conditions by analyzing human palms and nails using image matching technique*, Int. J. Sci. Engin. Res. **6** (2015), no. 10, 609–613.

[2] P.H. Bhapkar, T.Y. Puttewar and R.Y. Patil, *Topic name-nail lacquers in nail diseases*, IOSR J. Pharmacy **3** (2013), no. 9, Pp 24–48.

[3] P. Hardik and D. Shah, *The model of nail color analysis: An application of digital image processing*, Int. J. Adv. Res. Comput. Sci. Software Engin. **3** (2013), no. 5.

[4] S. Hsiung Lee and C. Sing Yang, *Fingernail analysis management system using microscopy sensor and blockchain technology*, Int. J. Distrib. Sensor Networks **14** (2018), no. 3, 1550147718767044.

[5] T.S. Indi and Y.A. Gunge, *Early stage disease diagnosis system using human nail image processing*, Int. J. Inf. Technol. Comput. Sci. **7** (2016), 30–35.

[6] T.S. Indi and S.D. Raut, *Person identification based on multi-biometric characteristics*, Emerg. Trends Comput. Commun. Nanotech. Int. Conf., 2013, p. 45–52.

[7] S. Kaur, M. Sandhu and J. Kaur, *Analysis of various image segmentation techniques using MATLAB*, Int. J. Adv. Res. Comput. Sci. Software Engin. 6 (2016), no. 3, 836–841.

[8] F. Noriaki and H. Kiyoshi, *Fingernail detection system using differences of the distribution of the nail-color pixels*, J. Adv. Comput. Intell. Intell. Inf. **17** (2013), no. 5.

[9] G. Sahoo and Y. Kumar, *Analysis of parametric and non parametric classifiers for classification technique using WEKA*, I. J. Inf. Technol. Comput. Sci. **4** (2012), no. 7, 43–49.

[10] V. Sharma and A. Shrivastava, *System for disease detection by analyzing finger nails color and texture*, Int. J. Adv. Engin. Res. Sci. 2 (2015), no. 10, 1–6.

[11] G. Singh, N. Sadath Haneef and A. Uday, *Nail changes and disorders among the elderly*, Indian J. Dermatol. Venereol. Lep. **71** (2005), no. 6, 386–392.

[12] A. Tsuda, H. Kato and A. Yoneyama, *Nail art simulation by hue difference feature with light*, Proc. 2009 IEICE Gen. Conf., Japan, 4 March, 2009, p. 290.

[13] A. Tsuda, S. Ueno and S. Sakazawa, *A method for extraction of nail area under varying luminance for the 3D nail art system on mobile phones*, IEICE Tech. Committee. PRMU **110** (2011), 381-–364.

[14] M. Waseem Khan, *A survey: Image segmentation techniques*, Int. J. Future Comput. Commun. **3** (2014), no. 2, 89–93.

[15] N.M. Zaitoun and M.J. Aqel, *A survey on image segmentation techniques*, Int. Conf. Commun. Manag. Inf. Technol. Procedia Comput. Sci. **65** (2015), 797–806.

[16] Z. Markov and I. Russell *An introduction to the WEKA data mining system*, ACM SIGCSE Bull. 38 (2006), no. 3, 367–368.

[17] https://www.dailymail.co.uk/health/article-15289/How-nails-reveal-health.html

[18] http://mrishaanshareef.blogspot.com/2008/10/finger-nails-to-predict-health.html

[19] https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/nails-fingernail-and-toenail-problems