JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

An Improved K-means based Nevus and Melanoma Classification System from Non-Dermoscopy Images using Deep Learning

Siddharth Kumar
M. Tech (ECE
Swami Vivekanand Institute of Engineering & Technology, Banur (Punjab), India,

Abstract: Over the last decades, digital image processing-based skin lesion segmentation and classification have been improving steadily to provide a more accurate classification results in the area of medical science. Segmentation and classification of the skin lesion from the non dermoscopic images is very challenging task due to the complex structural properties of the images and need improvisation in the existing work by utilization of improvement in segmentation and feature selection approach to select on optimal feature according to the nevus and melanoma type of skin cancer. A research based on the combination of various algorithms for nevus and melanoma classification has been presented in this paper. In this research, we develop an improved K-means based Nevus and Melanoma Classification (NMC) system from non-dermoscopy images using Deep Learning is proposed. The concept of Speed up Robust Feature (SURF) along with the hybridization of Artificial Bee colony (ABC) and Cuckoo Search Algorithm (CSA) for feature selection being used.
Here, Convolutional Neural Network (CNN) is used as an Artificial Intelligence (AI) technique with that helps to train the NMC system. By utilizing the concept of improved K-means, the accuracy of skin lesion segmentation is increases in terms of accuracy. So, classification efficiency of the NMC model is also improved as compare to existing works. At last, the performance of the NMC system is calculated to validate the model and this shows that it is possible to use optimized SURF as a feature extraction technique in order to classify the nevus and melanoma skin cancer with minimum error rate and the simulation results clearly show the effectiveness of proposed NMC system.

Index Terms **- Nevus and Melanoma Classification (NMC) System, SURF Descriptor, ABC, CSA, Artificial Intelligence (AI) and CNN**

I. INTRODUCTION

In the past few decades, the frequency of nevus and melanoma skin cancer has been increasing [1]. Where, a nevus is a type of benign skin cancer that is detected using the human skin lesion and cause due to the proliferation of cells that produce pigments which is known as melanocytes. While melanoma is a type of malignant skin cancer in the form of aggressive development that can develop from a benign type skin lesion [2]. Basically, melanoma is also originating from the same type of pigment-producing nevus cells in human body. Both type of research such as clinical and statistical have confirmed that nevus and melanoma is rapidly evolving, causing high mortality among the population worldwide.

In 2012, it was estimated that 76,250 cases of invasive melanoma had diagnosed in the United States, and an estimated 9,180 cases resulted in death [3]. Australia has one of the highest incidences of skin cancer in the world and more than 1,890 people in Australia had died from skin cancer per annum [4]. The growth rate of skin cancer is shown in Fig. 1.

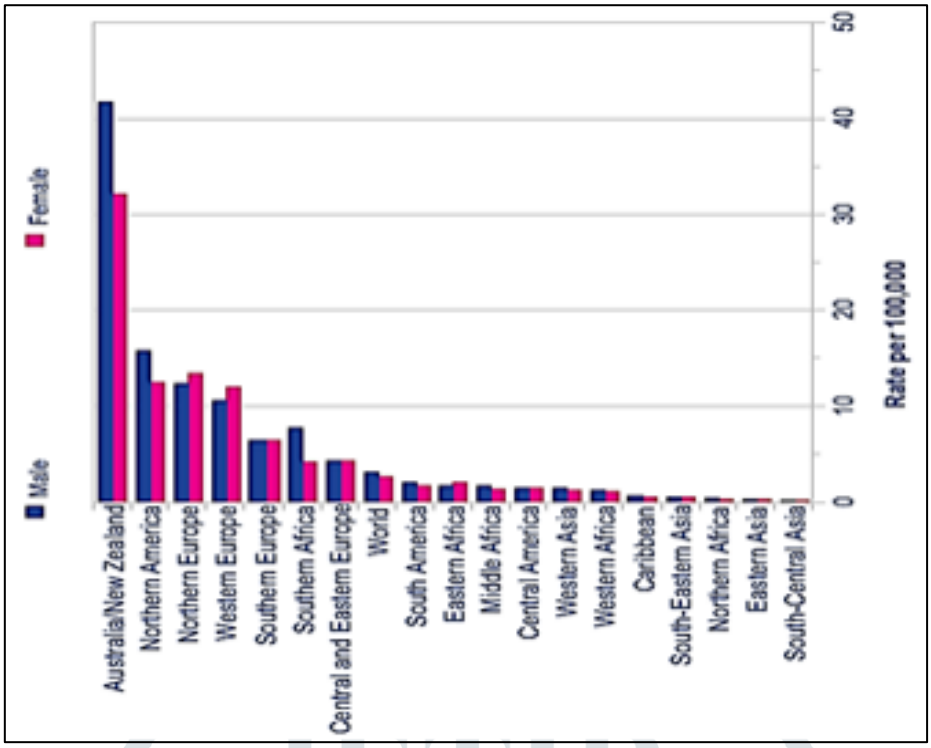


Fig. 1. Rate of Skin Cancer Rate Worldwide

Nevus and melanoma skin cancer has the ability to deeply invasive and the most hazardous feature of melanoma is that it spreads widely throughout the body by lymphatic and blood vessels. Therefore, the early diagnosis of melanoma is the main factor in the diagnosis of the disease [5]. Early detection of skin cancer is still essential in solving public health from both malignant melanoma and benign nevus skin cancer. To solve this type of problem, we present an improved K-means based Nevus and Melanoma Classification (NMC) System from non-dermoscopy images using Deep Learning. Main purpose to develop a NMC system is to solve out the segmentation and classification problem of skin cancer from the non-dermoscopy images. The block diagram of the proposed NMC system is shown in the Fig. 2.

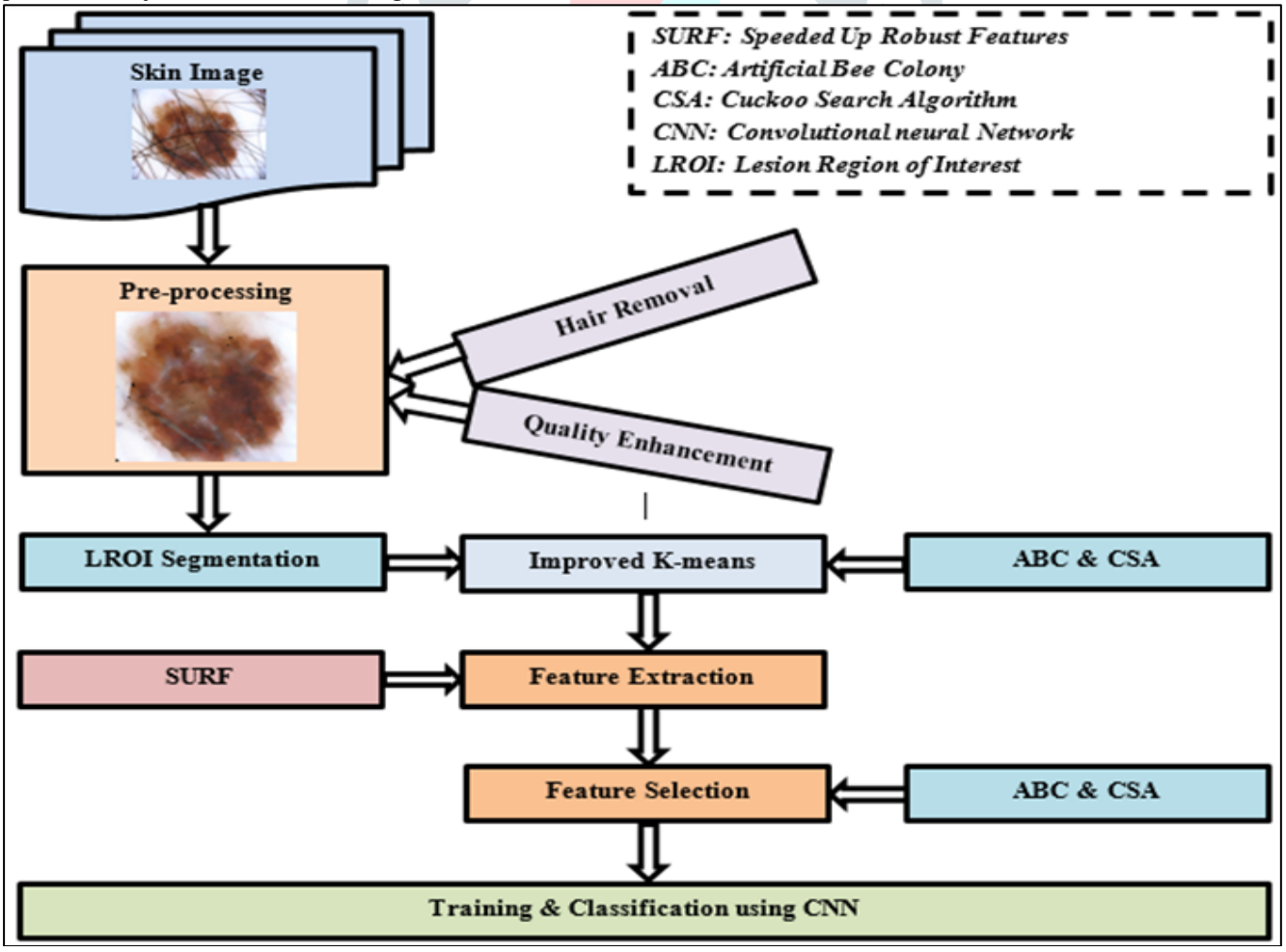


Fig. 2. Block Diagram of NMC System

The block diagram of the proposed NMC system is shown in the above Fig. 2 that depicts the entire procedure of implementation of NMC system. Image processing and computer vision based NMC system generally make use of four characteristic steps:

- Removal of hair over skin lesion

- ❖ Detect boundary of lesion from skin images for Lesion Region of Interest (LROI) segmentation
- ❖ Calculate high level features using SURF
- ❖ Select only appropriate feature sets using swarm-based meta heuristic algorithm

The proposed NMC system offer several foremost advantages over those traditional or machinery-based skin cancer detection system. Some advantages are given as:

- ☞ Main advantage is the designing cost of digital image processing-based NMC system is cheaper and the easy to implement as compare to the other NMC system.
- ☞ The response time of NMC system to detect the skin lesion and classify the diseases from image is faster as compare to any other traditional skin lesion-based skin cancer detection and classification systems.
- ☞ Developed NMC system has the ability to operate on large dataset of non-dermoscopic images.

Basically, in this research article, we focus on the development of an NMC system using the concept of the digital image processing algorithm with the optimized SURF by hybridization of two different swarm algorithms named as ABC and CSA. Here, the concept of Convolutional Neural Network (CNN) is used as classifier to train the NMC system by using dataset of non-dermoscopic images and the significant contributions in this research are as following:

- ➢ We present a brief survey about the existing Nevus and Melanoma Classification models to find out the inferences drawn.
- ➢ Pre-processing steps are used to segment the LROI from the skin image for better feature analysis using improved K-means with hybrid swarm-based technique (ABC and CSA).
- ➢ To extract feature from LROI, SURF descriptor is used with hybrid feature selection mechanism as a feature optimization or selection approach.
- ➢ To detect and classify the skin cancer from non dermoscopic images, CNN is used as an artificial intelligence technique.
- ➢ To validate the NMC model, comparison with existing state of arts is done on the basis of the parameters such as Accuracy, Sensitivity, F-measure, Precision, Mathew Correlation Coefficient (MCC), Dice, Jaccard and Specificity.

The rest of the research article is organized as: Section 2 portrays about the related literature works undergone in the recent past about the nevus and melanoma classification. Section 3 addresses the method and material of proposed NMC framework. Section 4 talks about the simulation results and finally, a strong conclusion is given to the current research work in Section 5 with future direction.

II. RELATED WORK

In this section, we present a brief survey relate to the nevus and melanoma classification from skin non dermoscopic images to find out the related issues in the existing models. In 2020, F. A. Damian et al. had conducted research on feature selection of non-dermoscopic skin lesion images for nevus and melanoma classification. Basically, authors used two datasets with 70 melanomas and 100 nevi were investigated and they used fast Fourier transform (FFT) to normalize the amplitude of features that is computed as a Fourier transform descriptor and focuses on geometric signatures of skin lesions using the frequency domain information. The receiver operating characteristic (ROC) curve and area under the curve (AUC) were employed to ascertain the relevance of the selected features and their capability to differentiate between nevi and melanoma. Here, AI (asymmetry index) features are circuited after the Otsu's method is used to segment the lesion region optimally by using grayscale image into a binary image conversion by setting threshold T values to minimize the overlapping of the class distribution. The asymmetry index and eccentricity, together with F6 Hu's invariant moment, were fairly competent in providing a good separation between malignant melanoma and benign lesions. From the simulation results analysis, 85% accuracy is observed and need to improvisation of medical science by utilizing the better pre-processing as well as segmentation techniques [6]. Siegel et al. in 2019 discussed about the ratio of cancer by conducting the survey, according to the American Cancer Society, the amount of new cancer instances and deaths resulting from cancer and the latest cancer information collected. In the United States 1,762,450 fresh instances of cancer and 606,880 fatalities from cancer are recorded. The general cancer death rate decreased by 27 percent from 1991 to 2016 or 2,629,200 lower cancer fatalities. As recorded in the past, the incidence rate of cancer in 2006-2015 was stable in females and decreased by 2% per annum in males, and the cancer death rate reduced by 1.4 percent and 1.8 percent per year in 2007-2016 [7]. Also in 2019, Paulson et al. proposed a system to detect the immunotherapy for skin cancer, throughout all the tumor type infections; skin cancers are higher sensitive to immune system and to validate clearly immunologic prevention of skin cancer in some patients, mostly cells are developed may be too early or late depends on the immunotherapy. This paper mainly focuses on the existing methodology with immunotherapy for skin cancers to date, and also the present observation of the methodology of resistance to immunotherapy [8]. In 2019, O'Sullivan et al. have proposed to predict the risk of skin cancer by using inner tanning devices that allowed Canada to forecast the possibility and amount of skin cancers dependent on tanning in Canada in 2015 and to differentiate between age and sex. It has been clarified that UV-releases tanning tools that induce melanoma and skin cancer that is not melanoma. The main purpose of this approach is to increase the use of UV protection chemicals and UV-Health regulators must take strength [9]. Yazawa et al. in 2019 explored the impact of the various elements on the dragline's physicochemical and mechanical characteristics. Upon removal of skin constituents, the crystal structure and mechanical characteristics are not significantly changed, indicating that structural and mechanical characteristics are determined by the main region of dragline silk fibers. However, the outer layers have little effect on super contraction, suggesting that they are not serving as protection against water molecules penetration. The exterior coats, on the other side, give a few securities beside incorporation of protease [10]. Mukherjee et al. in 2019 examined the detection of malignant melanoma, a form of skin cancer, by Neural Network classifier. Particle Swarm Optimization (PSO), a common met heuristic method, is used to find the optimum amount of neurons in multi-layer wireless network (MLP) concealed structures. Using a total of 1875 color, texture and shape characteristics obtained from dataset 170 color pictures, with triple cross-validation using a two-layer neural network, an accuracy of 85.9 percent is accomplished, which is 4.9 percent greater than earlier recorded outcomes with the same dataset [11]. In 2019, Muhammad et al. provided premature identification of pancreatic cancer is difficult because cancer-specific diseases only happen at a developed level and there is a lack of a secure testing instrument for high-risk nurses. Artificial Neural Network (ANN) has been implemented to tackle early detection of skin. Trained and evaluated using information collected from the National Health Interview Survey (NHIS) and Pancreatic, Lung, Colorectal and Ovarian Cancer (PLCO) datasets of 800,114 participants, together with 898 individuals identified with pancreatic cancer [12]. Akar et al. in 2019 established a cloud-based diagnostic system for skin lesions

using neural convolution networks. The diagnosis is done by a two-stage CNN pipeline in which a preliminary CNN conducts quality controls on customer requests and a CNN diagnosis whose reliability is similar to the dermatologist's point. Yields probabilities over seven distinct types of lesions corresponding to the classifications used for practice in the ISIC 2018 dataset. Experimental findings combine with the ResNet with such a small, multi-stream system developed by scratch where both networks ' outputs are connected to a mixed final layer. After, combining network shows high accuracy and high generalization results [13]. In 2018, Drost et al. explained about the previous research work done for the model of cancer, such preclinical models are compulsory to translate cancer research more efficiently translation into newer therapy schemes for cancer patients. In a setting that reaches the behavioral atmosphere, genetic manipulation of organoids enables modelling of disease. They have evaluated organoid tumor protocols and how they can be used in cancer research as an alternative model [14]. Also in 2018, Didona et al. have concentrated on several non-melanoma skin cancer variables were analyzed in detail and concentrated on molecular and environmental factors. Non-melanoma skin diseases (NMSCs) are among the most common malignancy in the world, 99% of which are basal cell carcinomas (BCCs) and squamous skin carcinomas (SCCs). NMSCs are completely non-lethal and cosmetically curable, so they are not reportable in most cancer registries around the world, but are presently presenting an increasing global health issue due to the increased incidence. Keratinocytes are both basal cells and squamous neurons, so keratinocyte cancer is sometimes called BCC and SCC [1].

Basically, above-described literature survey focused on different solutions to provide better accuracy of proposed an improved swarm-based K-means for skin lesion segmentation and detection from dermoscopy images using deep learning. After analysis of existing research work, the following points are highlighted as an inference drowns.

- For skin cancer detection in existing work, the segmentation of skin lesion from dermoscopic images is one of the most important procedures and it should be better in medical data processing.
- But in the exiting work, basic threshold-based segmentation methods are adopted which cannot achieved entire lesion region and mixed the background and foreground data.
- Pre-processing is one of the most important steps for a classification or detection model and for cancer detection using skin lesion segmentation it should be more appropriate. In existing work, hair removal and dermoscopic image enhancement was proposed by authors but removal of hair is not proper for such cases like if contrast of image is low then information loss is high with enhancement procedures.
- The entire computation arena is focusing on developing the enhanced classification structure but the root of the classification is a well-arranged targeted feature label which can be achieved by optimization techniques. The optimized feature set will always produce good classification accuracy during the skin cancer classification.
- So due to lack of feature optimization approach in the existing work, the selection of feature set is not good and it is depended on the optimization techniques. Hence a cross validation of the optimization is found to be missing which can produce a better feature set according to the type of dermoscopic data.
- In previous work, contour-based segmentation is applied but no threshold is justified which leads to error prone segmentation. This proposes can be improved and its way is justified in the problem statement.
- Swarm-based optimization is used for the optimization of the segmented region with improved fitness function but it is a density-based optimization algorithm which always needs large volume of data to be processed. So, validation of swarm-based techniques is an important step which is missing in existing work. Need to validation the other swarm-based algorithms like Cuckoo, Firefly, Whale etc. also be tried which works with low density as well.

So, in this research, we try to solve the existing problem by utilizing the concept of SURF as feature descriptor along with the hybrid feature selection mechanism using ABC and CSA with CNN to train and classify the nevus and melanoma skin cancer.

III. METHODS & MATERIALS

In this research article, we proposed NMC system using the concept of improved K-means for LROI segmentation and then optimized SURF is used to train and classify the nevus and melanoma skin cancer from non-dermoscopic image. Meanwhile all the implementation of the proposed NMC system need to cover the important steps of implementation. The entire methodology of proposed NMC system is dived into two different sections named as Training and Testing. The brief details about both phases are given in the below section of this paper and Fig. 3 show the flowchart of the proposed NMC system. The procedural and working steps of proposed NMC system is described in this section of research article according to the above shown flowchart in the Fig. 3. We use different step like pre-processing, SURF feature extraction and selection using hybrid swarm-based Meta heuristics technique with novel fitness function and CNN as classifier is used to train the NMC system. The subsequent steps demonstrate the variety of phases that need to be accomplished in the development of proposed NMC system and the used steps are given as:

Step 1. To design a model using the concept of GUI for simulation of proposed an improved swarm-based K-means for skin lesion segmentation and detection from non-dermoscopy images using deep learning.

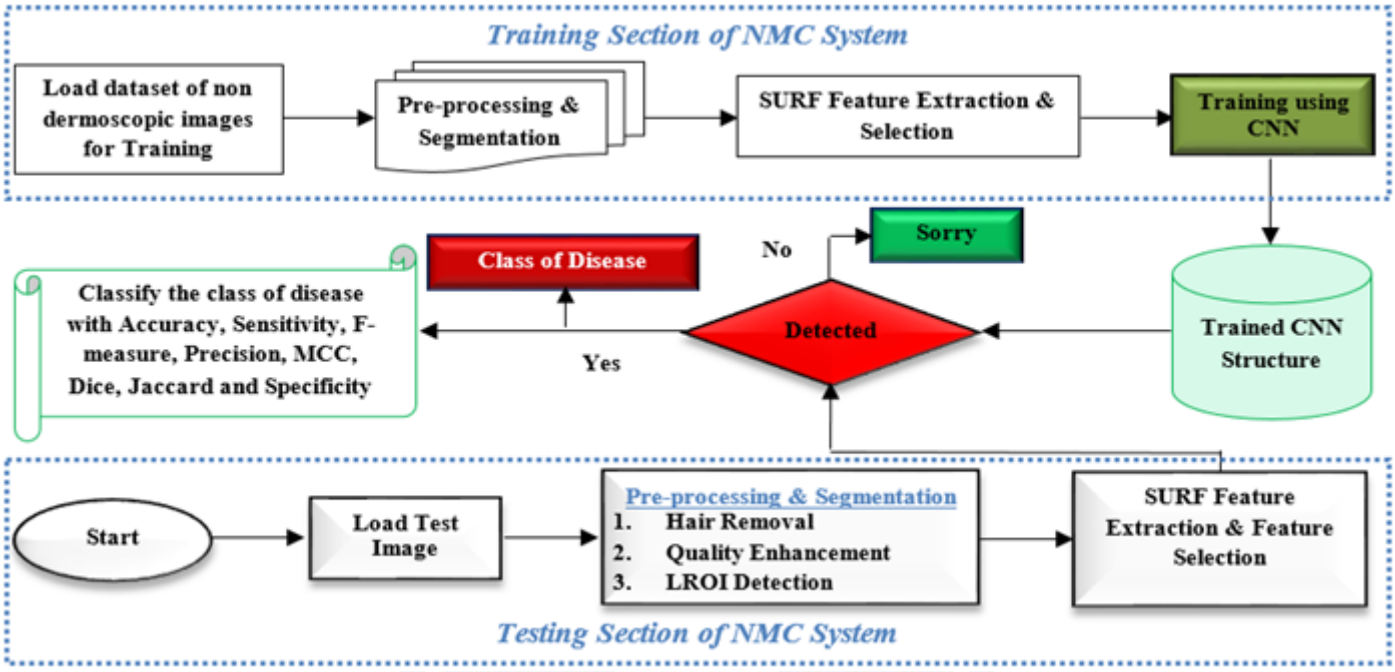


Fig. 3. Block Diagram of NMC System

Step 2. Upload test/train non-dermoscopy images for the simulation of proposed NMC model for the training as well classification purpose. The algorithm and flow of data uploading is written as:

Algorithm 1st: Dataset Uploading	
	N_T←No. of Images
Input	Folder←Location of dataset folder to load images
Output	Image Data ← Uploaded Skin Images
Steps	***Start Uploading***
1.	N_T = Number of *images* in Dataset F_{COUNT} = 0 // Frame count
2.	**For I in range of N_T**
3.	[File, Path] = Browse (Folder, N_T)
4.	Full Address = String concatenate (Path, File)
5.	Image Data [I] = Image read (Full Address)
6.	**End – For**
7.	**Return:** Image Data
8.	**End – Algorithm**

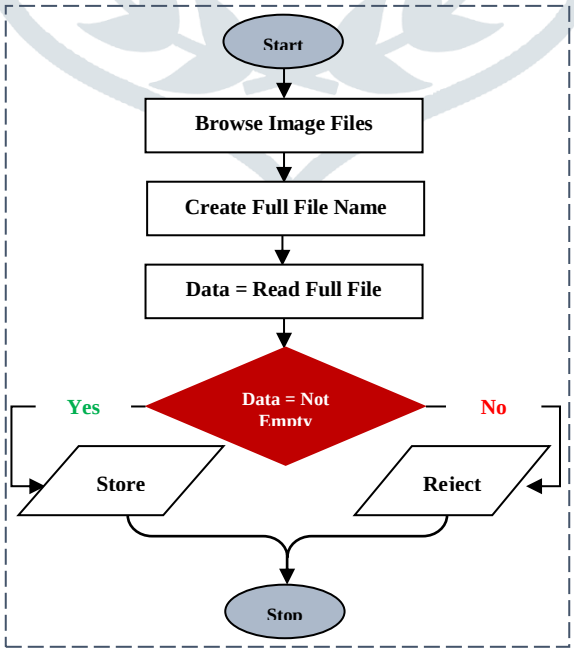


Fig. 4. Flow of Algorithm 1

Step 3. Apply pre-processing on the uploaded test dermoscopy images to segment the LROI using an improved K-means clustering technique with hybrid swarm technique. Firstly, pre-processing is performed to enhance the quality of dermoscopic images and remove the various type of noise that are inherited in the skin lesion images in proposed system. The algorithm of pre-processing for is written as:

Algorithm 2nd: Pre-processing on Skin Lesion	
Input	N_T←No. of Images Image Data ← Uploaded Skin Images
Output	P-Data ← Pre-processed Data
Steps	***Start Pre-processing***
1.	Set I_{COUNT} = 0 // Image count
2.	**While has Image Count (N_T)**
3.	I_{COUNT} = I_{COUNT} + 1
4.	P-Data = HR (Image Data)
5.	P-Data = Enhancement (P-Data)
6.	Save to P-Data with I_{COUNT} in Folder using location
7.	**End – While**
8.	**Return:** P-Data
9.	**End – Algorithm**

Here, the concept of hair removal is used to clean the skin lesion image suing the Hair Removal (HR) Algorithm that is written as:

Algorithm 3rd: HR	
Input	Image Data ← Uploaded Skin Images
Output	C-Data ← Clean Data
Steps	***Start Hair Removal***
1.	[Height, Width, Plane] = Size of Image Data
2.	Define, Patch Size = 15 and Pad Size = 7
3.	Creat Pmat = Pad array (Image Data, [Pad Size, Pad Size]) // It is blank mask according to the Pad Size
4.	**For j = 1 → Height**
5.	**For i = 1 → Width**
6.	Patch = Pmat(j: (j + PatchSize-1),i:(i + patchSize-1), All)
7.	**End – For**
8.	**End – For**
9.	A= atmLight (double (I), Zmat, Plane) // Apply Atmospheric Light Adjustment
10.	C-Data = zeros (size (I)) // Blank matrix with size of I
11.	**For ind = 1 → Plane**
12.	C-Data(:,: ,ind) = A(ind) + (I(:,:,ind) - A(ind))./max(Oimg,0.1) Where, 0.1 is lightning coefficient
13.	**End – For**
14.	**Apply Morphological Operations:** Structure Elements = strel ('disk', 3) Closeimg = imclose(C-Data, Structure Elements) Image Error = double (Closeimg) - double (Gimg) Dilated Image = (Image Error>5)
15.	C-Data = Dilated Image
16.	**Return:** C-Data
17.	**End – Algorithm**

Step 4. To design an accurate skin lesion segment and classification model, selection of popper lesion region is the major factor and need to remove extra part form the images that is called background. In this step, firstly, we perform hair removal approach to clear lesion region from hair and then intensity based image quality enhancement is used to improve the specific pixel points of a hair free dermoscopy images.

Step 5. After pre-processing on skin lesion images, segmentation process in performed to segment the LROI from dermoscopic image. Here, traditional K-means segmentation technique is used to segment the LROI in terms of fore front class and discard the rest data from the image that are assumed as backdrop. But, K-means algorithm faced the mixing problem of pixel for front class and backdrop data, so, the concept of the hybrid swarm-based optimization (Artificial Bee Colony (ABC) and Cuckoo Search Algorithm (CSA)) technique will be used due to their searching ability. The algorithm of improved K-means segmentation is written as:

Algorithm 4th: Improved K-means	
Input	P-Data ← Pre-processed Data
Output	LROI ← Lesion Region of Interest
Steps	***Start Segmentation***
1.	Set I_{COUNT} = 0 // Image count
2.	[R, C, P] = size (P-Data)
3.	P-Data = double (P-Data)
4.	Number of Part, N_P = 2

5.	Index = K-means(P-Data, N_P)
6.	Label Image = reshape(Index, R, C)
7.	Data Position = find (Label Image >0)
8.	Data = Label Image (Data Position)
9.	**Initialize ABC & CSA parameter** – Iterations (T) – Population Size (S) – Lower Bound (LB) – Upper Bound (UB) – Fitness function
10.	Calculate, T = Size (P-Data)
11.	Define fitness of Combined Fitness function using ABC: $f(fit) = \begin{cases} 1 & if pixel is less \\ 0 & otherwise \end{cases}$
12.	**For I = 1 → T**
13.	$fs = \sum_{i=1}^{P} Data(i)$
14.	$ft = \frac{\sum_{i=1}^{P} Data(i)}{Length of feature}$
15.	$Threshold_{value} = CSA(P, T, LB, UB, N, f(fit))$
16.	**End – For**
17.	Threshold = $Threshold_{value}$
18.	Mask = Morphological (Index, Threshold)
19.	Boundaries = bwboundaries (Mask)
20.	Segmented Region = Boundaries
21.	LROI = SL Image × Segmented Region
22.	**Return:** LROI
23.	**End – Algorithm**

Step 6. After ROL segmentation, we need to extract the feature pattern based on their pixel pattern using the SURF descriptor. Here, we select the SURF descriptor as a feature pattern extraction approach due to stability and invariance nature of features and SURF return more appropriate feature set for segmented ROL. The SURF descriptor is a fast and robust algorithm that helps to extract the local, invariant and oriented feature set from the ROL of dermoscopic images.

Algorithm 5th: SURF Descriptor	
Input	LROI ← Lesion Region of Interest
Output	F-points ← SURF feature points
Steps	***Start Feature Extraction***
1.	Calculate size, [R, C] = size (LROI)
2.	**For m = 1→ R**
3.	**For n = 1→ C**
4.	Scaling = LROI (m, n, 8) // Scaling of LROI into 8 X 8
5.	Ext_Det = Extrema (Scaling(m, n)) // detect extremes of scaled frame
6.	F-points = Localization (Ext_Det (m, n))
7.	Check: F-points orientation = F-points vary or not
8.	**If Variation occurs**
9.	Discard F-points
10.	**Else**
11.	F-Points = F-points (m, n)
12.	**End – If**
13.	**End – For**
14.	**End – For**
15.	**Return:** F-points
16.	**End – Algorithm**

Step 7. To achieve better classification accuracy of proposed model, this step is performed in terms of the selection of best feature set from the high dimensional feature data that is returned by the SURF descriptor. Because numerous features data are presents in the SURF feature and it should be considered as irrelevant data and do not involve in the training scenario because they increase the chances of error in the model. So, for selection of appropriate feature pattern, hybrid swarm-based optimization is used with a novel fitness function and the algorithm is written as:

Algorithm 6th: Feature Selection using hybridization of ABC & CSA	
Input	F-points ← SURF feature points
	MAX_{ITR} ←Maximum iteration
	Pop← Number of populations
	FF←Fitness Function
Output	OF-points ← Optimized feature points
Steps	***Start Feature Selection***
1.	Set FT_{COUNT} = 0 // Feature count
2.	**Call FF using the ABC for CSA,** $F(f) = \begin{cases} 1\ (True); & if\ F_s \geq F_t \\ 0\ (False); & Otherwise \end{cases}$ (1)
3.	**While ITR = MAX_{IRT}**
4.	$FT_{COUNT} = FT_{COUNT} + 1$
5.	F_S = F-points (ITR)
6.	F_t = mean (F-points)
7.	OF-points = IWD (F_s, F_t, Pop, FF)
8.	**End – While**
9.	**Return:** OF-points as a set of optimized feature points
10.	**End – Algorithm**

Step 8. Using the above mention hybrid swarm-based optimization, select only appropriate set of feature according to the skin cancer classes with fitness criteria. After the feature pattern selection, used these feature as input of CNN classifier to train the proposed skin lesion segment and cancer classification model and here, the pattern net based CNN as a classifier or deep learning approach is used and the algorithm of CNN is written as:

Algorithm 7th: CNN for NMC System	
Input	OF-points ← Optimized feature points
	Group ← Nevus and Melanoma
	N ← Number of Neurons
Output	NMC-Structure→ Trained structure with Detection Results of NMC System
Steps	***Start Training***
1.	**Initialize the basic parameters of CNN** –Number of Epochs (E) –Number of Neurons (N) –Number of Layers (L) –Performance: Cross entropy –Activation Techniques: ReLU – Data Division: Random
2.	**For i = 1 → OF-points**
3.	**If OF-points ε Nevus**
	G (1) = Features (OF-points)
4.	**Else if OF-points ε Melanoma**
5.	G (2) = Features (OF-points)
6.	**Else // Extra**
7.	G (3) = Features (OF-points)
8.	**End – If**
9.	**End – For**
10.	Initialized the CNN, **Net = Pattern-net (OF-points, G, N)**
11.	**Net = Train (Net, OF-points, G)**
Classification of Nevus & Melanoma	
12.	Test Frame = Simulate (Net, OF-point of Test)
13.	**If Matched with Nevus & Melanoma**
14.	Mark as cancer
15.	**Else**
16.	Normal
17.	**End – If**
18.	Calculate Parameters, P
19.	NMC-Structure = [Net, P]
20.	**Return:** NMC-Structure
21.	**End – Algorithm**

Step 9. Here, CNN as a classifier is used to train the NMC model based on the different skin lesion non dermoscopic image dataset. So, optimized set of feature pattern is considered as an input set of CNN with different skin cancer types such Melanoma or

Nevus from Basal Cell Carcinoma (BCC), Squamous cell carcinoma (SCC), Markel cell carcinoma (MCC), Cutaneous T-cell lymphoma and Kaposi sarcoma.

Step 10. At last of simulation, the performance parameters of propose NMC system is calculated in terms of Accuracy, Sensitivity, F-measure, Precision, MCC, Dice, Jaccard and Specificity. Based on the quantities parameters such as True Positive (TP), False Positive (FP), True Negatives (TN) and False Negative (FN), we evaluate the effectiveness of the proposed NMC system and the explanation of parameters are given as:

Precision: This parameter is used to calculate the efficiency of a classifier along with the proposed model. If the value of precision is high it means that the false positive rate is less and vice versa.

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP+FP})} \quad (2)$$

Where, TP→ It is the collection of all relevant testing feature according to the output
FP→ It is the collection of all irrelevant testing feature according to the output
TN→ It is the collection of all relevant training feature according to the output
FN→ It is the collection of all irrelevant training feature according to the output

Sensitivity or Recall: This term is used to measure the comprehensiveness of a classifier. More the value of recall indicates the less false negatives but improving recall usually decreasing the precision value.

$$\boldsymbol{Sensitivity\ or\ Recall} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP+FN})} \quad (3)$$

F-measure: It is the rate that is obtained by combining both precision and recall value and obtaining harmonic mean.

$$\boldsymbol{F-measure} = \boldsymbol{2} \times \frac{\boldsymbol{TP}}{\boldsymbol{2 \times TP + FP + FN}}$$

Or

$$\boldsymbol{F-measure} = \boldsymbol{2} \times \frac{\boldsymbol{Precison \times Recall}}{\boldsymbol{Precison + Recall}} \quad (4)$$

Accuracy: It is defined as the sentiments classified correctly with respect to the entire available classified sentiments.

$$\boldsymbol{Accuracy} = \frac{(\boldsymbol{TP+TN})}{(\boldsymbol{FN+FP+TP+TN})} \quad (5)$$

Error: It is the reverse of accuracy and calculated using given formula

$$\boldsymbol{100 - Accuracy = Error\ rate} \quad (6)$$

MCC: It is used in machine learning as a measure of the quality of binary (two-class) classifications, introduced by biochemist Brian W. Matthews in 1975.

$$\boldsymbol{MCC} = \frac{(\boldsymbol{TP \times TN - FP \times FN})}{\sqrt{((\boldsymbol{TP+FP}) \times (\boldsymbol{TP+FN}) \times (\boldsymbol{TN+FP}) \times (\boldsymbol{TN+FN}))}} \quad (7)$$

DC: It is also known as the Sørensen–Dice index and it is a statistical tool which measures the similarity between two sets of data. Formula of DC is written as:

$$\boldsymbol{DC} = \frac{\boldsymbol{2 \times TP}}{(\boldsymbol{2 \times TP+FP+FN})} \quad (8)$$

JC: It is defined as the rate of DC with respect to the 2 minus DC and also defined as the size of the intersection divided by the size of the union of two label sets, is used to compare set of predicted labels for a sample to the corresponding set of labels in TP.

$$\boldsymbol{Jaccard} = \frac{\boldsymbol{DC}}{(\boldsymbol{2-DC})} \quad (9)$$

Specificity: It is used to measures the proportion of TN that are correctly identified and the formula is Witten as:

$$\boldsymbol{Specificity} = \frac{\boldsymbol{TN}}{(\boldsymbol{TN+FP})} \quad (10)$$

Execution Time: The simulation time required to test the sentiments during experiment is known as computation time.

The simulation results analysis of the proposed NMC system is discussed in the next section of this research article with the comparison to validate the system efficiency.

IV. RESULTS AND ANALYSIS

In this section of research article, we describe the simulation results of the proposed NMC system to segment and classify the nevus and melanoma skin cancer from the non dermoscopic images.

Non Dermoscopic Skin Lesion Dataset: It contains the human lesion analysis towards nevus and melanoma detection and the dataset is in the form of non dermoscopic images. To capture images, the non dermoscopy process is used that is an imaging technique to eliminate the surface reflection of human skin. It provides improved diagnostic accuracy and the sample of non dermoscopic skin lesion dataset shown in the Fig. 5.

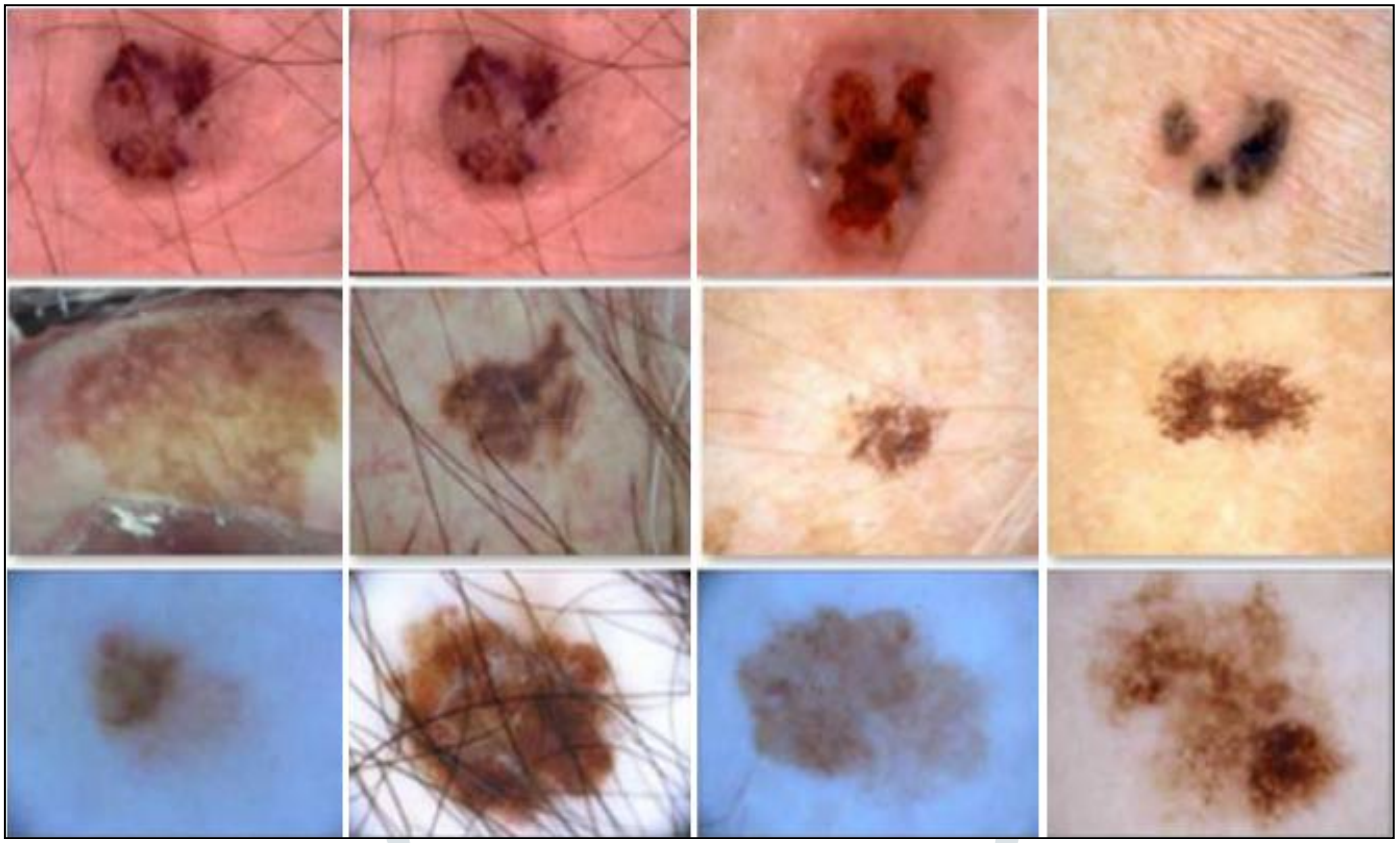
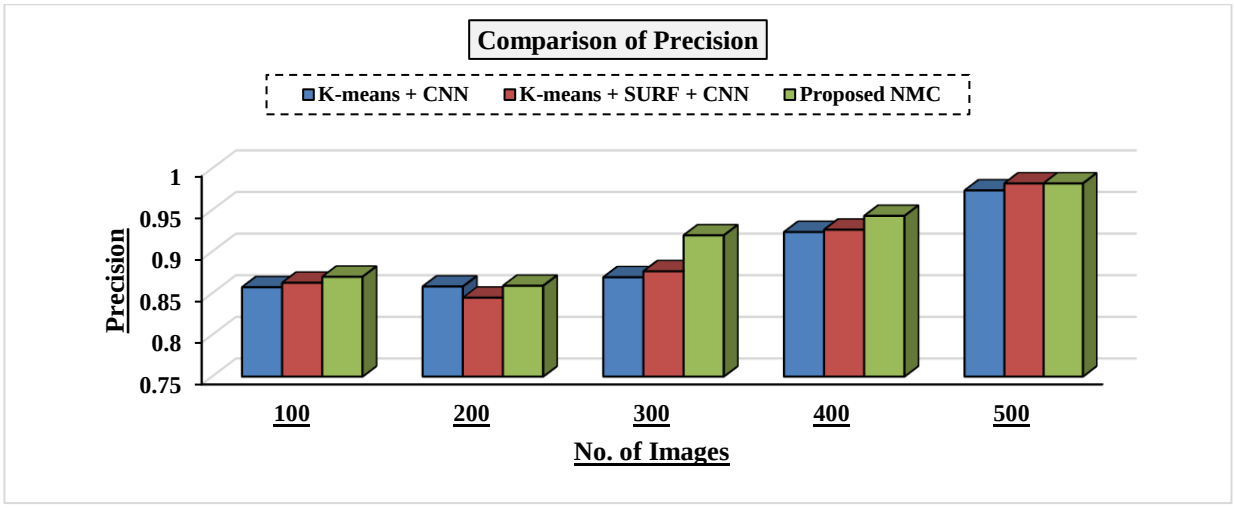


Fig. 5. Non Dermoscopic Skin Lesion Dataset

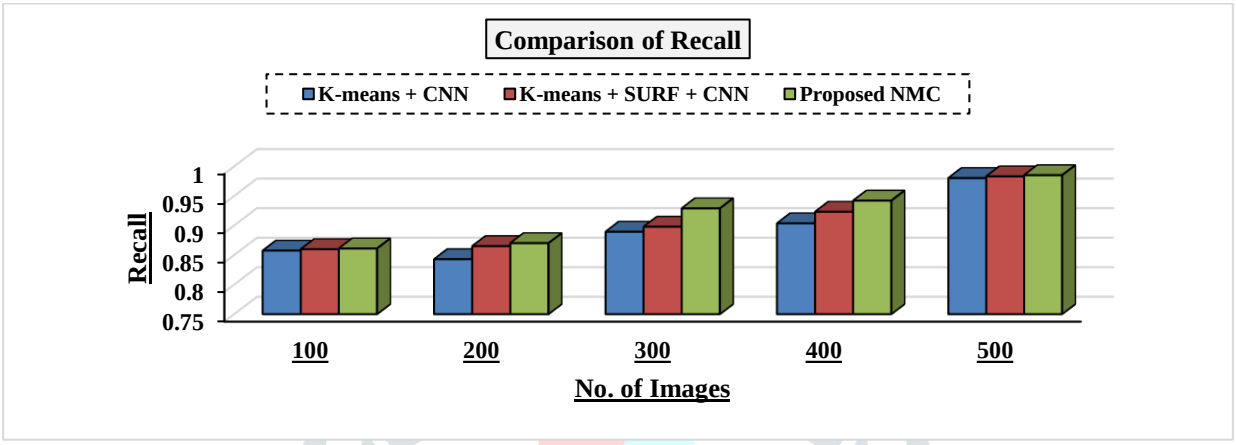
So, the proposed NMC system is test on the non dermoscopic image dataset of skin lesion. The simulation results of the proposed NMC system using the CNN are given in the Table I and their graphical representation is also described in the below section of the research article.

Table 1: Comparison of simulation results based on quantities parameters

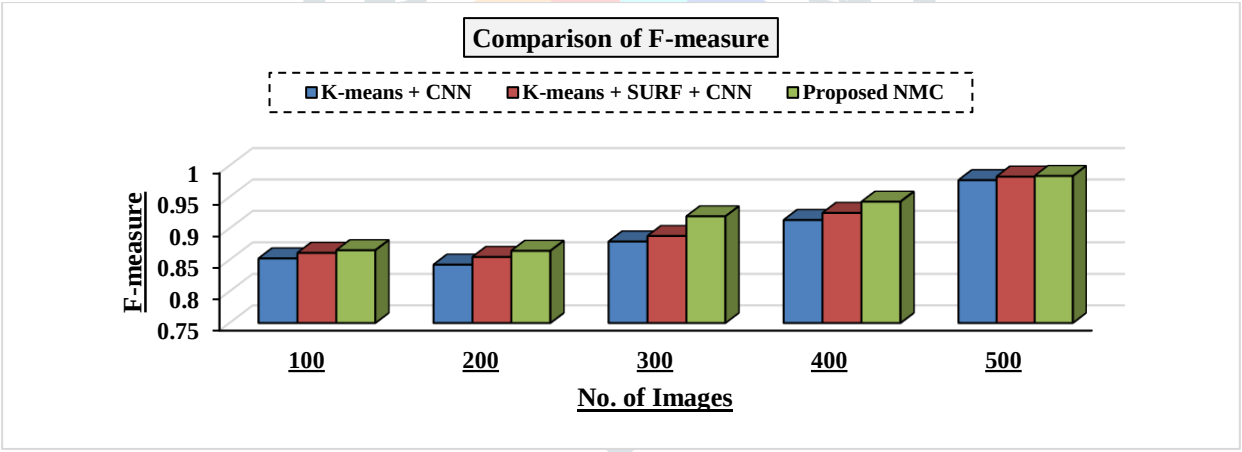
PARAMETERS		K-means + CNN		K-means + SURF + CNN	Proposed NMC
Precision	No. of Images	100	0.8574	0.8627	0.8699
		200	0.8582	0.8448	0.8591
		300	0.8694	0.8765	0.9197
		400	0.9237	0.9265	0.9431
		500	0.9737	0.9821	0.9821
Recall	No. of Images	100	0.8575	0.8599	0.8611
		200	0.8430	0.8651	0.8701
		300	0.8895	0.8979	0.9290
		400	0.9037	0.9232	0.9421
		500	0.9803	0.9831	0.9851
F-measure	No. of Images	100	0.8524	0.8612	0.8654
		200	0.8426	0.8548	0.8645
		300	0.8793	0.8880	0.9193
		400	0.9135	0.9248	0.9426
		500	0.9769	0.9825	0.9835
Accuracy	No. of Images	100	81.5816	87.8346	92.8432
		200	84.3495	89.8007	95.1030
		300	90.0117	90.1153	96.3699
		400	89.6862	94.7604	98.2569
		500	97.6035	98.5501	99.6902
Error	No. of Images	100	19.616	12.1654	7.1568
		200	17.4607	10.1993	4.8970
		300	10.9674	9.8847	3.6301
		400	10.3622	5.2396	1.7431
		500	3.8899	1.4499	0.3098


Comparison of Precision
K-means + CNN
K-means + SURF + CNN
Proposed NMC
Precision
1
0.95
0.9
0.85
0.8
0.75
100
200
300
400
500
No. of Images

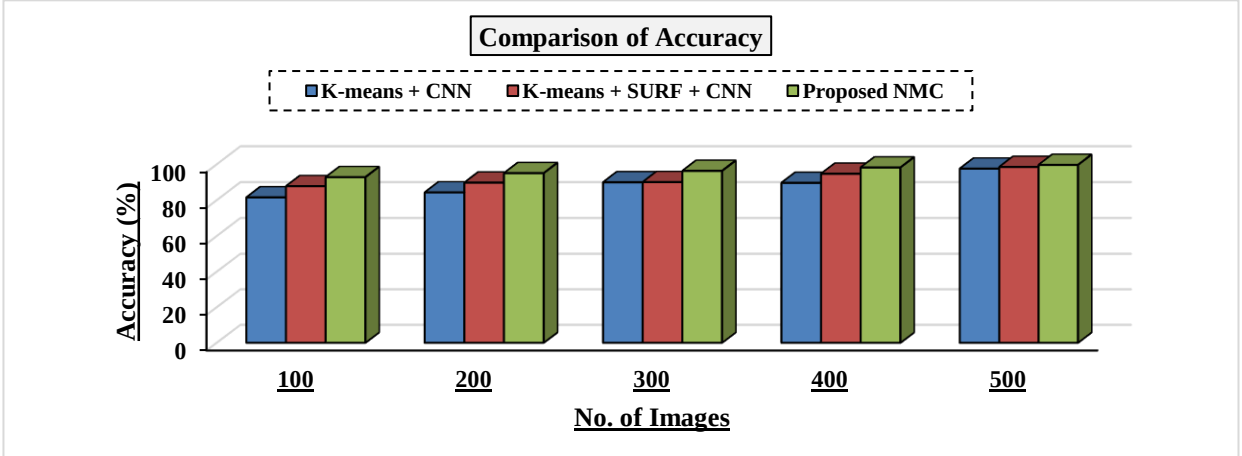
(a)


Comparison of Recall
K-means + CNN
K-means + SURF + CNN
Proposed NMC
Recall
1
0.95
0.9
0.85
0.8
0.75
100
200
300
400
500
No. of Images

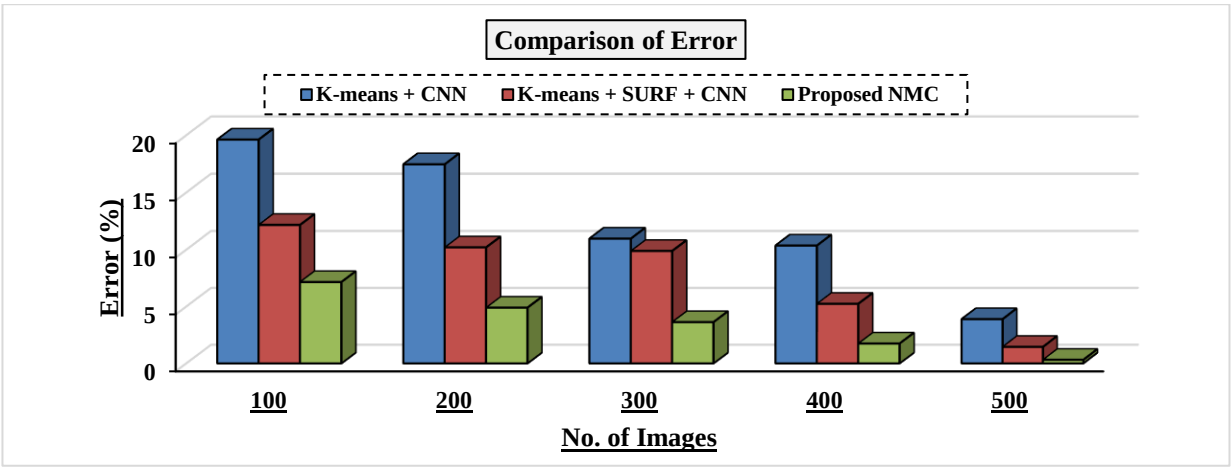
(b)


Comparison of F-measure
K-means + CNN
K-means + SURF + CNN
Proposed NMC
F-measure
1
0.95
0.9
0.85
0.8
0.75
100
200
300
400
500
No. of Images

(c)


Comparison of Accuracy
K-means + CNN
K-means + SURF + CNN
Proposed NMC
Accuracy (%)
100
80
60
40
20
0
100
200
300
400
500
No. of Images

(d)



(e)

Fig. 6. Comparison of simulation results based on quantities parameters (a) Precision (b) Recall (c) F-measure (d) Accuracy and (e) Error

From the Table 1 and Fig. 6, we observed that the simulation results of proposed NMC frameworks using the concept of hybridization of the K-means with ABC and CSA is superior to others modules in terms of the quantities parameters and improvements in quantities parameters are clearly visible in the Fig. 6 with average accuracy are 88.65%, 92.22% and 96.46% for K-means with CNN, K-means with SURF and CNN and Proposed NMC system respectively. So, we can say that the effect of proposed NMC model using CNN is far better than other combination. However, we need to validate the model based on similarities parameters such as MCC, DC, JC and computational time. Therefore, the simulation results comparative analysis based on the similarities value given in Table 2.

Table 2: Comparison of simulation results based on similarities parameters

PARAMETERS			K-means + CNN	K-means + SURF + CNN	Proposed NMC
MCC	No. of Images	100	0.8447	0.8893	0.9706
		200	0.8651	0.9157	0.9713
		300	0.8858	0.9394	0.984
		400	0.8882	0.945	0.9857
		500	0.9023	0.9485	0.9979
JC	No. of Images	100	0.8418	0.8982	0.9668
		200	0.8602	0.8993	0.9717
		300	0.8834	0.9165	0.9793
		400	0.8912	0.9363	0.9867
		500	0.8957	0.9489	0.9871
CD	No. of Images	100	0.8897	0.8941	0.9608
		200	0.8975	0.8986	0.9610
		300	0.9158	0.9076	0.9695
		400	0.9221	0.9215	0.9865
		500	0.9473	0.9272	0.9888
Time	No. of Images	100	8.25839	8.45225	8.5419
		200	8.37362	8.50933	8.53367
		300	8.79246	8.83669	9.18777
		400	9.01242	9.28575	9.4042
		500	9.74011	9.76711	9.79502

In terms of similarities parameters, also proposed NMC system is superior to others modules. However, to validate the efficiency of NMC system, we need to compare with state-of-the-art work based on their accuracy in Table 3.

Tbale 3: Comparison with existing works

Works	Accuracy (%)
F. A. Damian et al., 2020 [6]	83.50
FCM with PSO	87.54
K-means + CNN	88.65
K-means + SURF + CNN	92.22
Proposed NMC System	96.46

Graphically comparison of different frameworks such as K-means with CNN, K-means with SURF and CNN and proposed NMC system with existing work by F. A. Damian et al., 2020 [6] is shown in Fig. 7.

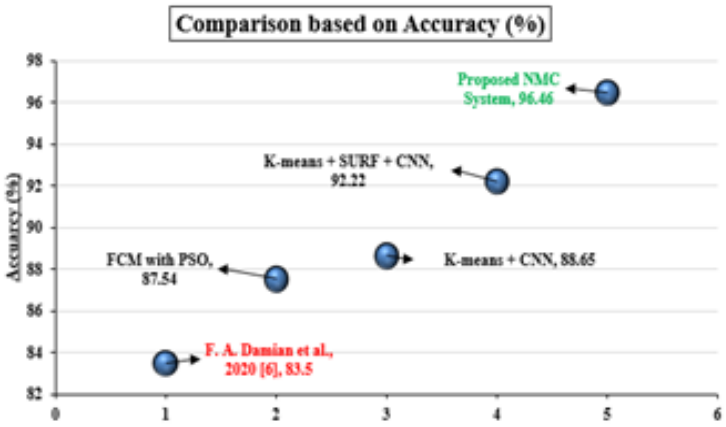


Fig. 7. Comparison of Proposed NMC System with Existing Work

Based on the comparative analysis of the existing work with proposed NMC system, we observed that the accuracy of the proposed NMC model is far better than the exiting work and it is possible by utilizing the concept of improved K-means for LROI segmentation and optimized SURF feature descriptor along with CNN and it show that the effectiveness of swarm algorithm hybridization for the feature selection.

V. Conclusion & Future Work

In this paper, an improved K-means based Nevus and Melanoma Classification (NMC) System from non-dermoscopy images using CNN as a Deep Learning is proposed. We have evaluated whether it is possible to use optimized SURF feature set in order to minimization of error rate in the proposed NMC system with hybridization of two different swarm algorithm named as ABC and CSA for segmentation as well as feature selection. We showed that, it is possible to make an efficient NMC system by utilizing the hybrid-based SURF feature along with the CNN as a classifier, which takes input as SURF feature sets to train the system and we trained the model based on the different images. From the simulation analysis of experimental results, we observed that the proposed NMC system is an innovative and successful step towards the nevus and melanoma classification. In future, proposed system could be designed behalf of the concept of real-time scenario.

References

[1] D. Didona, G. Paolino, U. Bottoni, C. C.- Biomedicines, and undefined 2018, "Non melanoma skin cancer pathogenesis overview," mdpi.com, doi: 10.3390/biomedicines6010006.

[2] F. Islami et al., "Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States," Wiley Online Libr., vol. 68, no. 1, pp. 31–54, Jan. 2018, doi: 10.3322/caac.21440.

[3] F. Navarro, … M. E.-V.-I. journal of, and undefined 2018, "Accurate segmentation and registration of skin lesion images to evaluate lesion change," ieeexplore.ieee.org, 2018, doi: 10.1109/JBHI.2018.2825251.

[4] S. Mustafa, A. K. A. I. T. (IWAIT), and undefined 2018, "A SVM-based diagnosis of melanoma using only useful image features," ieeexplore.ieee.org, Accessed: Jan. 19, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8369646/.

[5] L. Singh, R. Janghel, S. S.-P. C. Science, and undefined 2020, "Designing a retrieval-based diagnostic aid using effective features to classify skin lesion in dermoscopic images," Elsevier, Accessed: Jan. 19, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S187705092030733X.

[6] F. Damian, S. Moldovanu, N. Dey, A. Ashour, L. M.- Computation, and undefined 2020, "Feature selection of non-dermoscopic skin lesion images for nevus and melanoma classification," mdpi.com, Accessed: Jan. 19, 2022. [Online]. Available: https://www.mdpi.com/705114.

[7] C. Kaushal, A. Singla, P. P.-2021 9th International, and undefined 2021, "A Brief Review on Clustering Based Medical Image Segmentation Algorithms with Issues and Challenges," ieeexplore.ieee.org, Accessed: Jan. 19, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9596154/.

[8] K. Paulson, M. Lahman, … A. C.-I., and undefined 2019, "Immunotherapy for skin cancer," academic.oup.com, Accessed: Jan. 19, 2022. [Online]. Available: https://academic.oup.com/intimm/article-abstract/31/7/465/5310046.

[9] D. O'Sullivan, D. Brenner, … P. D.-C., and undefined 2019, "Indoor tanning and skin cancer in Canada: A meta-analysis and attributable burden estimation," Elsevier, Accessed: Jan. 19, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877782118306453.

[10] K. Yazawa et al., "Role of skin layers on mechanical properties and supercontraction of spider dragline silk fiber," Wiley Online Libr., vol. 19, no. 3, Mar. 2018, doi: 10.1002/mabi.201800220.

[11] S. Mukherjee, A. Adhikari, and M. Roy, "Malignant melanoma detection using multi-layer perceptron with optimized network parameter selection by PSO," Adv. Intell. Syst. Comput., vol. 812, pp. 101–109, 2019, doi: 10.1007/978-981-13-1540-4_11.

[12] W. Muhammad et al., "Pancreatic Cancer Prediction Through an Artificial Neural Network," Front. Artif. Intell., vol. 2, May 2019, doi: 10.3389/FRAI.2019.00002/FULL.

[13] E. Akar, O. Marques, W. A. Andrews, and B. Furht, "Cloud-Based Skin Lesion Diagnosis System Using Convolutional Neural Networks," Adv. Intell. Syst. Comput., vol. 997, pp. 982–1000, 2019, doi: 10.1007/978-3-030-22871-2_70.

[14] J. Drost, H. C.-N. R. Cancer, and undefined 2018, "Organoids in cancer research," nature.com, Accessed: Jan. 19, 2022. [Online]. Available: https://www.nature.com/articles/s41568-018-0007-6.