ISSN 2228-9860 eISSN 1906-9642 CODEN: ITJEA8

International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies

http://TuEngr.com

Skin Disease Stratification based on Regression with Fuzzy C-mean Adversarial Tunicate Algorithm (RFAT)

Ravinder Reddy Baireddy[1*], R.Nagaraja[2], G.Sudha[3*]

[1] *IT Department, Sreenidhi Institute of Science & Technology (SNIST), Hyderabad, INDIA.*
[2] *ISE Department, Bangalore Institute of Technology (BIT), Bangalore, INDIA.*
[3] *EEE Department, Bangalore Institute of Technology (BIT), Bangalore, INDIA.*
**Correspinding Authors (Email: ravibaireddy @gmail.com, profgsudha @gmail.com).*

Paper ID: 13A6D

Volume 13 Issue 6

Received 15 January 2022
Received in revised form 25 April 2022
Accepted 01 May 2022
Available online 07 May 2022

Keywords:
Cluster algorithm; Benchmark function; Stratification of skin disorder; Adversarial Tunicate Swarm Algorithm (ATSA); Mean Square Deviation (MSD); Dunn Index (DI); Tunicate Swarm Algorithm (TSA); Fuzzy clustering algorithm; En-ABC; Oppositional Tunicate Swarm Algorithm (OPTSA); ANN, LR-HID, ML-IDS; Synthetic dataset; NIB (IBRL); NSL-KDD dataset.

Abstract

Skin disease stratification is essential for individualized treatment. Stratifying skin disorders involves finding disease subgroups for effective therapy. Academics and the medical community have recently become interested in cluster algorithms for stratifying skin conditions. Cluster methods feature experimental sounds, a large number of dimensions, and insufficient interpretation. Cluster algorithms evaluate cluster quality with one internal operation. Creating a robust internal evaluation mechanism for all datasets is difficult. This research included several time-complexity investigations. Despite its prevalence, skin tone and hair color variations confound identification. Many studies cannot properly predict skin conditions. This paper presents an RFAT algorithm to overcome these difficulties. The image data used in this approach has been pre-processed, and Linear Regression is employed to remove unwanted material. The Adversarial Tunicate Swarm Algorithm partitions data using Fuzzy Clustering (FM) (ATSA). Optimal clustering-based data has been adapted to fulfill sub problems' ultimate purpose (A3). MSD and Dunn Index can help with this (DI). Logistic Regression can help choose attributes. Data classification uses Sigmoid and logistic regression. Clustered data output is divided by normal and aberrant. A performance investigation determined OPTSA's local optimum avoidance capacity. It has also been evaluated using ANN, LR-HID, ML-IDS, and En-ABC. Synthetic, NIB (IBRL), and NSL-KDD datasets are used for the proposed work's efficacy.

Disciplinary: Electrical Engineering, Medicine and Health Sciences.

Cite This Article:

Baireddy, R.R., Nagaraja, R., Sudha, G. (2022). Skin Disease Stratification based on Regression with Fuzzy C-mean Adversarial Tunicate Algorithm (RFAT). *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies, 13*(6), 13A6D, 1-15. http://TUENGR.COM/V13/13A6D.pdf DOI: 10.14456/ITJEMAST.2022.109

1 Introduction

With its promising findings, skin disease categorization and segmentation have been emerging in the artificial intelligence space. It is a standard procedure to employ clustering methods and support vector machines (SVMs) for segmenting and classifying skin diseases [1,2]. Overall, clustering algorithms are flexible, easy to implement, and capable of generalizing traits with a statistical variance that are similar. Fuzzy c-means, modified fuzzy c-means, and K-means were all used by Trabelsi et al. [3] to segment a skin ailment with an accuracy of about 83 percent in their study. They employed an ISODATA clustering technique to find the optimal threshold for skin lesion segmentation. Clustering a skin condition is limited by the availability of noise resiliency. Clustering methods rely on establishing a centroid in order to generalize a cluster of data. Outliers and noisy data impede the development of these algorithms. A fuzzy clustering model proposed by Keke et al.[4] is a more robust alternative to non-clustering strategies for dealing with noisy datasets, as opposed to the original fuzzy clustering methodology. High-dimensional data has made SVMs famous for their ability to decipher "...subtle patterns in noisy and complex datasets." Lu et al.[5] employed the radial basis kernel function to segregate erythema in the skin using SVMs that can differentiate nonlinear hyperplanes. An SVM-k-NN classifier combination by Sumithra et al. [6] was used for the segmentation and classification of skin lesions into five distinct categories. For classification, Maglogiannis et al. [7] utilized an SVM with a threshold on the RGB value. It is more difficult to extract features from data using SVMs than clustering techniques that are much more resilient. SVMs may potentially underachieve if they are not pre-processed such that hyperplanes may be clearly defined.

Because of the drawbacks of old techniques, convolution neural networks (CNNs) have become highly widespread because of their capacity to retrieve high-level data with minimal pre-processing. SVMs can reap the benefit of CNNs, such as resilience in noisy datasets without the requirement for efficient pre-processing, by absorbing picture context and retrieving high-level features by down-sampling. The pixels of a picture can be examined within the context of the picture itself rather than being examined as part of a dataset by CNNs. As a result, the quality of images is degraded when CNNs use down-sampling. Although the context is gained, the location of a target is lost as a result of downsampling. There are almost no issues with categorization; however, fragmentation is a little more complex because context and location are critical for the successful performance of the system. Up-sampling, which functions in the reverse direction of down-sampling, is required to resolve this problem, as it enhances the picture's resolution. A matrix is reduced to a feature map whereas a feature map can be increased to a larger matrix through up-sampling. The location of the targets can be segmented through learning to make a high-resolution image. So we apply an integration of up and downsampling, while solely down-sampling is used for categorization. When CNN models became too huge and sophisticated, skip-connections were proposed as a remedy to the degradation problem that develops. Both

segmentation and classification models use skip connections. To connect the down- and up-sampling portions in the segmentation model, blocks with similar feature numbers are used as connectors. These skip connections are represented in the classification model by inverted residual blocks. The complexity of our models can thus be increased without affecting their efficiency.

2 Data Clustering Exploiting FM and APTSA

The Adversarial Tunicate Swarm Algorithm is used in conjunction with the clustering technique of fuzzy C-mean to demonstrate how data might be clustered.

A. Adversarial Tunicate Swarm Algorithm (APTSA)

A Tunicate Swarm Algorithm (TSA) and Oppositional Learning Algorithm (OLA) combination are called Oppositional-based TSA (OBL). Following is more information on the Oppositional Tunicate Swarm Algorithm (OPTSA).

B. TSA (Tunicate Swarm Algorithm):

Using the tunicate, seafood can be ranked according to its quality. Moreover, the characteristics (Swarm intelligence incorporated with jet propulsion) of tunicate determines the optimal food source. Further, the jet propulsion must satisfy three criteria: i) close to the search agent ii) ignorance of divergence among the agents, and iii) movement toward the search agent. The characteristics of Swarm will be upgraded by the best optimal solution. The numerical expression is illustrated in the following section.

a. Ignorance of Divergence among the search agent: The divergence among the search agent can be denied by estimating a new location for the search agent.

$$\vec{D} = \frac{\overrightarrow{H}}{\overrightarrow{N}} \quad (1),$$

$$\overrightarrow{H} = b_2 + b_3 - \vec{F} \quad (2),$$

$$\vec{F} = 2.b_1 \quad (3).$$

$\vec{F}$ represents the water flow advection and $\overrightarrow{H}$ denotes the gravity force. ‘b1, b2, and b3’ are the arbitrary variables and lie from 0 to 1. Moreover, the social force between the search agent is determined by $\overrightarrow{N}$,

$$\overrightarrow{N} = [R_{\min} + b_1.R_{\max} - R_{\min}] \quad (4).$$

Equation (4.17), $R_{\min}$ and $R_{\max}$ determines the initial and subordinate speeds which are used to create social interaction. Let us consider the value of $R_{\min}$ is 1 and $R_{\max}$ is 4.

The movement to the best behavior direction: After ignoring the divergence of the neighbors, the search agents shift towards the best neighboring side.

$$\overrightarrow{RC} = \left| \overrightarrow{GL} - r_{andom} \cdot \overrightarrow{R_p(x)} \right| \tag{5}$$

Here, AS indicates the interval between the food source and search agent. Meanwhile, x represents the current execution of the tunicate and GL determines the location of the food source. $[r_{andom} \in 0,1]$ and it's an arbitrary constant $\overrightarrow{R_p(x)}$ that denotes the tunicate location.

Convergence rate: Convergence rate is the maintenance of convergence between the search agents. The upgraded position of $\overrightarrow{R_p(x)}$ the can is denoted as $\overrightarrow{SF}$.

$$\overrightarrow{R_p(x)} = \begin{cases} \overrightarrow{GL} + \vec{Y} \cdot \overrightarrow{AS} & if \quad r_{andom} \geq 0.5 \\ \overrightarrow{GL} - \vec{Y} \cdot \overrightarrow{AS} & if \quad r_{andom} < 0.5 \end{cases} \tag{6}$$

i. Characteristics of the swarm:

After completing the convergence rate the location of search agents is upgraded to replicate the swarm characteristics of the tunicate.

$$\overrightarrow{R_p(x+1)} = \frac{\overrightarrow{R_p(x)} + \overrightarrow{R_p(x+1)}}{2 + d_1} \tag{7}$$

ii. Adversarial-Based Learning (ABL):

Heuristic methods have been used to derive the optimal solution from arbitrary choices in an initial population in common. The algorithm's convergence time is determined using the obtained solution. Convergence times are shorter if the initial solution is nearer ideal; otherwise, they are longer. Opposition learning also enhances the initial population [6]. The executions were repeated until the enhanced solutions were obtained. Moreover, the OBL is used to choose the estimation of the current candidate solution, and hence optimal threshold learning and searching were utilized to stimulate the OBL process. The following section clarifies the OBL steps.

Opposition Number: The opposition integer $\bar{o}$ for the real number (o) can be determined as $\bar{o} = m + n - o$ (i.e. $(o \in [m, n])$). The multidimensional behavior of the opposition integer can be determined as above.

Opposition point: Assume the opposition point $P = (x_1, x_2, \ldots, x_D)$ in D-dimensional space and $o \in [m_a, n_a]$, $y_n \in R$. Here, $n = \{1, 2, 3, \ldots, D\}$. The opposition point can be expressed numerically as

$$o = m_a + n_a - o_a \tag{8}$$

Optimization in terms of Opposition: Assume the candidate solution as $P = (x_1, x_2, \ldots, x_D)$ in D-dimensional space. Moreover, the fitness function $\zeta(P)$ can be estimated by considering the

opposition point of P $(\overline{P})$. Hence the fitness function for the opposition point must be greater than the point P, i.e., $\zeta(\overline{P}) \geq \zeta(P)$. Thus the optimization has been obtained for the opposition point.

iii. Adversarial Tunicate Swarm Algorithm (ATSA)

The local optimization problem can be solved by utilizing ATSA. To attain these, two functions have to be considered i.e., Mean square deviation (MSDδ), and Dunn Index (DIμ). Of these MSDδ should be minimized and DIμ should be maximized. This can be explained below.

- The tunicate population can be initiated by utilizing OBL in Equation (2)
- Choose the initial parameters for the greatest aggregation of executions
- Estimate the fitness value for each search agent by utilizing Equation (3)
- In the present space search, determine the most effective search agent and calculate the optimal value for every agent in the space.
- The up-gradation of each search agent location can be estimated by Equation (4)
- Estimate the fitness value for the upgraded search agent and if it is greater than the already existing value then upgrade the solution as R_p. Repeat the executions until the optimal solution has been attained, i.e., R_{pB}.

C. Fuzzy Clustering Algorithm

A fuzzy clustering algorithm is exploited to partition the clusters into O clusters. Moreover, each data points hold its membership grades and lies between 0 and 1. The data points which are nearer to the cluster exhibit a maximum membership grade than the other data points. Consider, $D_\eta = \{d_j\}$ as the dataset along with $j = \{1,2,...,m\}$ and $O = \{q_i\}$ be the center cluster, and $i = \{1, 2,3,...,k\}$. Besides FCM based clustering is carried out to perform the reduction of the objective function I_n.

$$I_n = \sum_{j=1}^{m}\sum_{i=1}^{k} P_{ji}^{n} \left| d_j - q_i \right|^2, \qquad where\ 1 \leq P < \infty \tag{8}$$

Here m and k are real values that are higher than unity and the data point $P_{ji} \in [0,1]$ distributes the membership grades. Moreover, in q_i the cluster the d_j is used to denote the data point which holds membership grades. Further, Euclidean distance is evaluated to analyse the similarity between the data points d_j and their respective cluster q_i. Further, the upgraded membership grade P_{ji} for the cluster centre q_i can be numerically expressed as,

$$P_{ji} = \sum_{l=1}^{k} \left(\frac{\left| d_j - q_i \right|}{\left| d_j - q_i \right|} \right)^{-2/(n-1}} \tag{9}$$

However, the cluster center for the iteration *i* can be expressed as

$$O_i = \frac{\sum_{j=1}^{m} P_{ji}^{n} d_j}{\sum_{j=1}^{m} P_{ji}^{n}} \quad (10).$$

Algorithm 1: FCM clustering model

Input: OPTSA's optimal value (RPB)and the data points to be clustered

Set $k \leftarrow R_{pB}[]$ and terminal criterion $\Re$

for

Each j and fixed n

do

Initialization of P_{ji} as the fuzzy membership grade / $\sum_{i=1}^{k} P_{ji} = 1$ and $P_{ji}^{n} \in [0,1]$

end for

initialization of P with ρP

O_i can be calculated with the equation (5)

The Euclidean distance δ_{ji}^{n} can be estimated between the i^{th}

Use equation (8) to calculate the Euclidean distance δ_{jin} between the i^{th} cluster centre and data object $j^{th} \in D_s$

if

$\delta_{jin} > 0$

then

Use Equation (9) to update P_{ji}^{n}

else if

$\delta_{jin} = 0$

then

Cluster j with i the cluster data point

$\sum_{i=1}^{k} P_{ji}^{n} = 1$

if

$\wp P < \varsigma$

then

Again use equation (6) to determine the C_i

else

end if

Output: Set of k clusters

In every clustering algorithm evaluation of the complexity of C (M, n) is predominantly important to avoid complexity. However, the proposed ATSA provides better optimal output for FM. Moreover, the process followed is explained in

Algorithm 2. The computational complexity evaluated by using the proposed work can be determined as $C((R_p)+(M.n))$. RP denotes the aggregations of solution whereas, n and M represent the number of data objects and are constant.

Algorithm 2: ATSA based FM Algorithm

Input: Set the Solution population of OPTSA with the highest values of iterations, OPTSA parameters, and testing set $D_{\eta,test}$

Initialize the tunicate population $(\overrightarrow{R_p})$ and candidate solution $P=(x_1,x_2,\ldots,x_D)$

The fitness value can be estimated by exploiting Equation (7)

Initialize $R_{p,\,bezt}[]\leftarrow null$

The best behaviour of the swarm can be determined by using Equation (8)

The convergence value of the best agent can be evaluated by using Equation (9)

for

if

$(r_{andom}\leq 0.5)$

then

$swarm\leftarrow swarm+\overrightarrow{SF}+\vec{B}\cdot\overrightarrow{RC}$

else

$swarm\leftarrow swarm+\overrightarrow{SF}--\vec{B}\cdot\overrightarrow{RC}$

end if

end for

if

$F(\overline{P})\geq F(P)$

then

Replace P with $(\overline{P})$

else

update the solution $R_{p,\,bezt}[]\leftarrow R_p$

end if

set $R_{p,\,bezt}[]\rightarrow$ cluster centroids of the optimal position for FCM

Call (Algorithm 1)

Output: Custer centroids k with optimal positions

2.1 Cluster Handling Stage

To manage the development of clusters, it is necessary to do clustering data handling after data collection is complete. This has been discussed in more detail below.

(a) *Cluster Expansion*

The expansion is started by setting a cluster head or cluster center. The objects having better similarities are considered as cluster centers. Typically, cluster extension is used to fix the issues of scattered clustering whereas the cluster formulation is being done. This problem can be answered using the sum of squared errors (SqE). Consider the kth cluster with b_k^p cluster members.

Where $p = \{1,2,3,...,n\}$ $k = \{1,2,3,\ldots,x\}$.

Here x- the total number of members in the kth clusters.

The sum of the squared error of the multidimensional vector d_v can be estimated by

$$\forall k_i;\ S_qE = \sum_{k=1}^{x}(d_v - \hat{d})^2 \qquad (11),$$

where $k_l; l = \{1,2,3,...,k\}$ - The entire number of clusters that have been formed by deploying FCM.

Moreover, the dispersion problems can be removed by reassigning clusters by ignoring and insisting on the unsuitable objects in other clusters.

Cluster Integration

The cluster Integration has been carried out by taking minimum values of S_qE that the objects which are nearer to the centroids.

$$k_e = \frac{1}{x}\sum_{k=1}^{x} d_k^p \qquad (12).$$

The cluster centers are allocated based on the clusters and for clusters *g* and *h* the cluster centers are denoted as d_e^g and d_e^h. Its variance is denoted as σ^g and σ^h. The constraints used to satisfy two clusters can be given as d^g and d^h [9].

$$\left|d_e^g - d_e^h\right| \le \frac{1}{2}(\sigma^g + \sigma^h) \qquad (13).$$

The total S_qE evaluated for all the k clusters can be given as (Montgomery *et al.* (2009)),

$$S_qE_{total} = S_qE_{k^1} + S_qE_{k^2} + \ldots + S_qE_{k^k} \qquad (14).$$

Otherwise, $$S_qE_{total} = \sum_{l=1}^{k}\sum_{k=1}^{x}(d_k^l - d)^2 \qquad (15).$$

The squared error for clusters 1 and 2 are given as $S_qE_{k^1}$ and $S_qE_{k^2}$. To integrate the clusters, the clusters with the least impact S_qE_{total} have been chosen.

3 2.2 Intrusion Detection Stage

Among the key divisions of intrusion detection are point intrusion detection and collective intrusion detection, both of which are discussed in greater detail below.

(a) *Point Intrusion Detection*

The Euclidean distance can be used to calculate the distance between the centre object (ke) and each data object and can be written as

$$\delta_{Fe} = \sqrt{\sum_{p=1}^{n} (d_F^p - k_e^p)^2} \tag{16}$$

The cluster radius (Rc) is determined by the distance between the data object(d_c)and center of cluster(k_e). Hence the distortion measurements can be determined as

$$R_c = \max(\delta_{ce}) \tag{17}$$

Here,

$$\delta_{ce} = \sqrt{\sum_{p=1}^{n} (d_c^p - k_e^p)^2} \tag{18}$$

Here, if $\delta_{ce} \le R_c$ then the d_F belongs to $\forall d_j$ cluster.

(b) *Collective Intrusion detection*

Collective intrusion detection can be performed by estimating the distance between the clusters. The interval between the clusters d^g and d^h from the cluster k^c can be determined as,

$$\delta_{c/g,h} = \frac{(m_g + m_c)\delta_{gc} + (m_h + m_c)\delta_{hc} - m_c\delta_{gh}}{m_g + m_h + m_c} \tag{19}$$

The distance between the object d^g and d^h from the cluster k^c can be indicated as $\delta_{c/g,h}$. Moreover, $m_g, m_h, and\ m_c$ represents the objects g, h, and c from the cluster. However, the distance can be calculated as δ_{gc} for the clustering object d^g and d^c. δ_{ac} is the distance between the d^c and d^a cluster. Similarly, δ_{ba} is the distance between the d^a and d^c cluster

4 Experimental Analysis

The performance of the implemented technique can be analyzed by exploiting some of the performance metrics. Then the resultant outcomes are compared with the state-of-art methods. To performance, analysis is carried out in execution with Cloudsim. These are explained below.

4.1 Details of the Dataset

To evaluate the performance, the NIB, IBRL, and NSL-KDD datasets, as well as the synthetic dataset, were used. Table 1 contained all of the information. In this data set, we consider skin disease images.

Table 1: Dataset description.

Parameters	Datasets				Image type
	Synthetic	*NIB*	*IBRL*	*NSL-KDD*	
# Events	71,002	3,56,242	3.1 million	1,28,448	Dermoscopic images
# Classes	5	2	2	4	Clinical images
# Features	8	2	8	10	Clinical images

Synthetic Dataset

The synthetic dataset is produced by the controlled testbed which was tested in a cloud computing environment. The taken testbed exhibit four hosts and is interlinked with the Local Area Network. They are listed below. (i) Offering web services and giving support for a variety of Virtual Machines, the Apache HTTP server has been implemented Exploiting this service also initiates the movement and traffic. Thus, it can be used to insist on malicious traffic for the detection of intrusion. (ii) The hardware emulation can be achieved by exploiting QEMU9, (iii) the virtualization support for the host can be provided by enabling KVM8, and (iv) Libcap 11 has been used to capture the packet traces among the different nodes. Thus, the Malicious traffic and benign plays a vital role in network traffic formation.

Numenta Intrusion Benchmark (NIB):

The National Institute of Standards and Technology (NIST) dataset is the most widely used and available dataset for intrusion detection analysis. A mixture of fake and real-world datasets is included in this collection.

Intel Berkeley Research Lab (IBRL) dataset

The Intel Berkeley research lab dataset contains time-stamped information that was grasped from the 54 Mica2Dot sensors. This also gathers information about the temperature, voltage, humidity, and light via small DB in the network query processing systems.

NSL-KDD dataset

NSL-KDD is another and updated version of KDD 99. Moreover, KDD 99 holds redundancies and hence this selected dataset can be used to overcome them. This dataset can also be exploited to identify the intrusion in the big data of the military. Usually, it exhibits 24 different types of attacks, and further, it is enclosed in four different classes. They are Probe attack, R2L
(Remote to Local), DOS (Denial of Service) and U2R(User to Root).

Image analysis relies largely on the following two criteria in order to succeed. Many other clinical indicators can be used to make a proper diagnosis, including asymmetry and irregularity of the lesion's border Additionally, consider the importance of precisely identifying borderlines when it comes to extracting additional clinically significant data, such as strange dots and colour variation. Drawing on a dermoscopic image for inspiration (Figure 1),

4.2 Measure of Work

The performance of the proposed intrusion detection algorithm can be analyzed by using different performance metrics such as rate of false positive, true positive, and accuracy value.

$$A_c = \frac{TN + TP}{FN + FP + TN + TP} \quad (20),$$

$$FPR = \frac{FP}{FN + FP} \quad (21),$$

$$TPR = \frac{TP}{FN + TP} \quad (22).$$

Here, TN is used to denote the true positive, and TN is used to indicate the true negative rates. Similarly, FP and FN are used to determine the rate of false negative and false positive.

5 Performance Evaluation of TSA

TSA is used not only for detecting the hyper-plane but also for distinguishing between labelled samples/data and unlabeled samples/data using the classifier TSA. Our TSA classifier, which makes use of feature extraction, is quite simple in its recognition of the presence of disease on the skin. The following is an example of the objective function of TSA:

A comparison of the proposed work's effectiveness to other similar ones has been created to assess its overall quality. These include Gray wolf Optimizer (Srivastava et al. (2019) [8], Ensemble Artificial Bee Colony (En-ABC) [9], and Tunicate Swarm Algorithm (TSA).

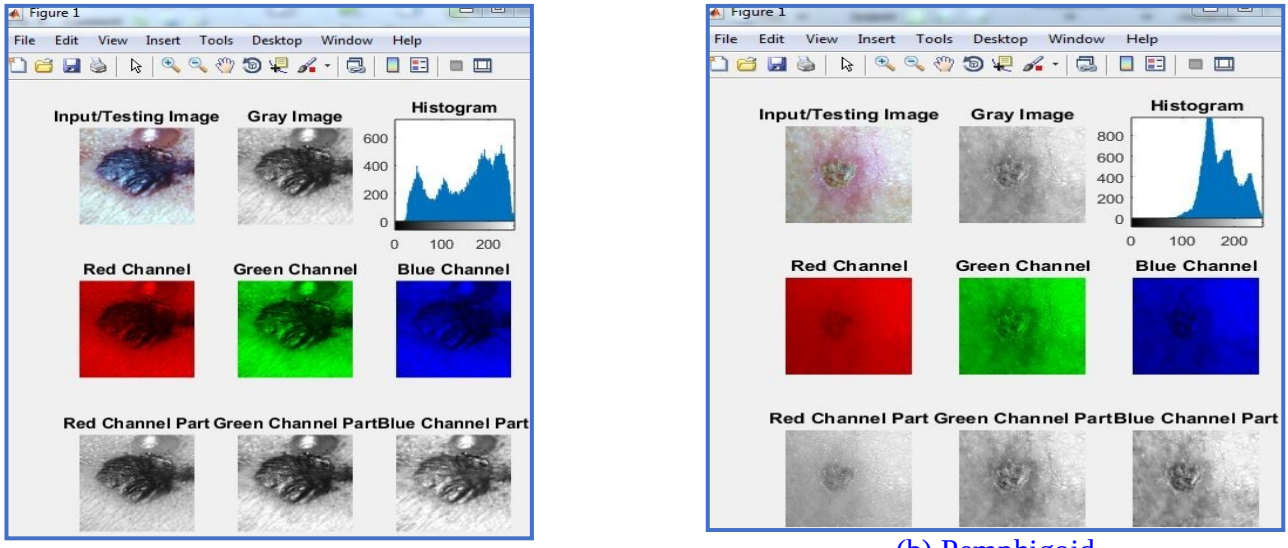
Algorithm 1: TSA training process algorithm

1. Initialize parameter T and T^*
2. Apply inductive SVM on training data
3. Set positive label samples number
4. Find label samples for all unlabelled samples using decision boundary function value in where label samples are positive for the highest value of decision boundary function
5. Set T^*_{tmp}
6. Retrain SVM over the whole samples of labelled and unlabelled samples
7. Shift unlabelled samples to labelled samples and make the value of objective function
8. If all the data satisfy the shifting condition, then
   $T^*_{tmp}=T^*_{tmp}+1$, else go to step 7
   If $T^*_{tmp}>T^*$
9. then obtain the output data, else go to step 6
10. Stop

Testing and comparison are two separate processes that take place once an image feature has been determined. By comparing the new data with the TSA algorithm's training set, we can get data that can be categorized.

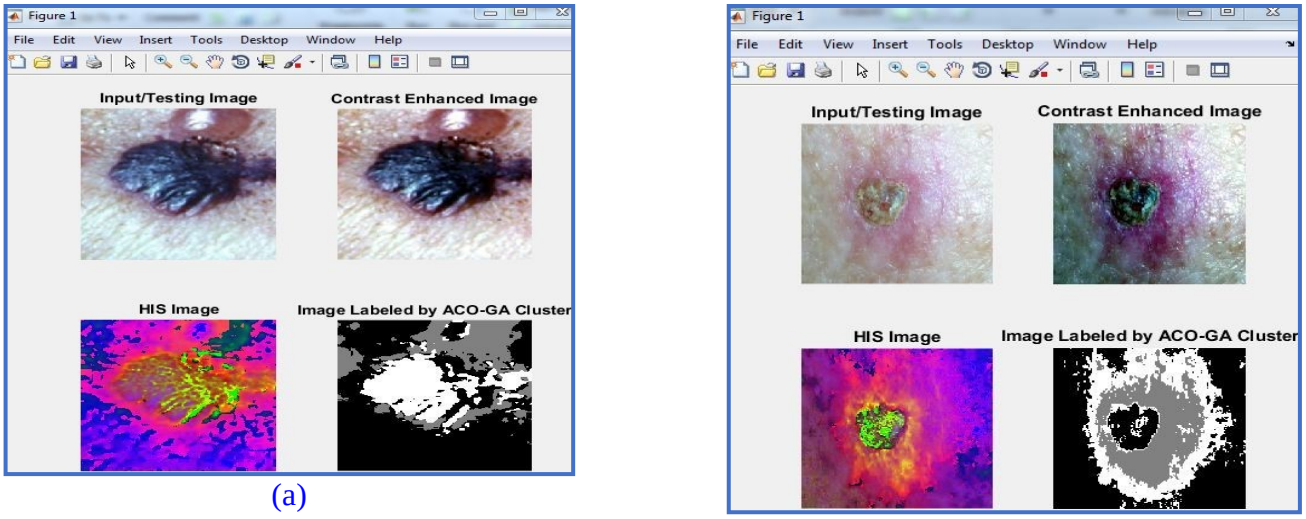
The TSA classifier is used to identify and detect skin diseases.

6 Results Analysis and Discussion

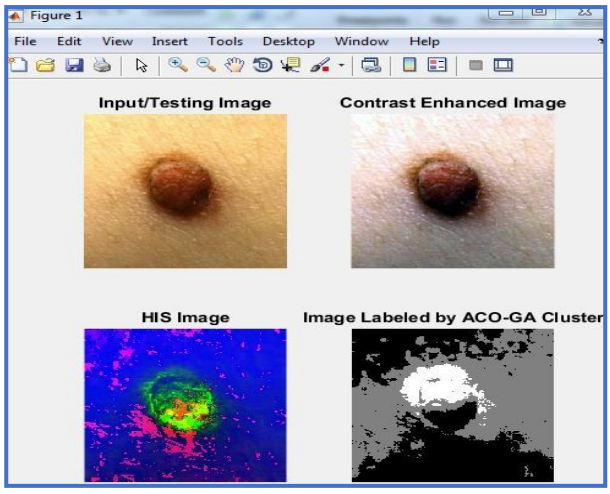


(a) Basal cell (b) Pemphigoid

Figure 1: Color transformation process of diseased skin mages.



(a) (b)

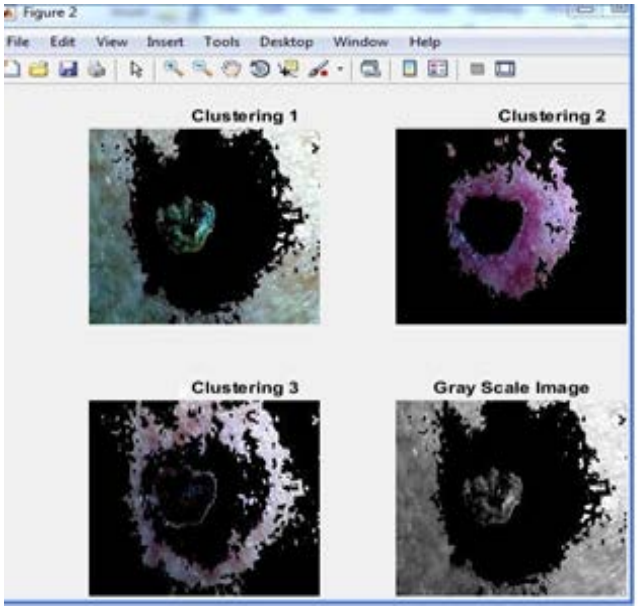


(c)

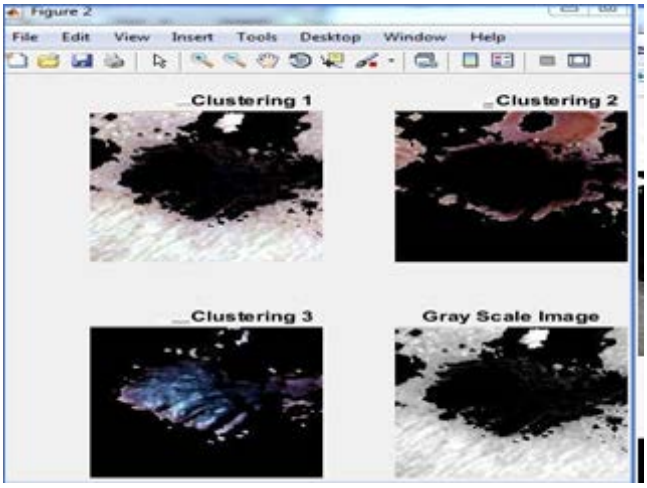
Figure 2: Testing infected skin image transformation to HIS, contrast-enhanced, and ACO-GA clustering labeled image.

MATLAB was utilized to construct a prototype for our unique hybrid intelligent RFAT technique, which is capable of accurately segmenting dermatological pictures. Support vector

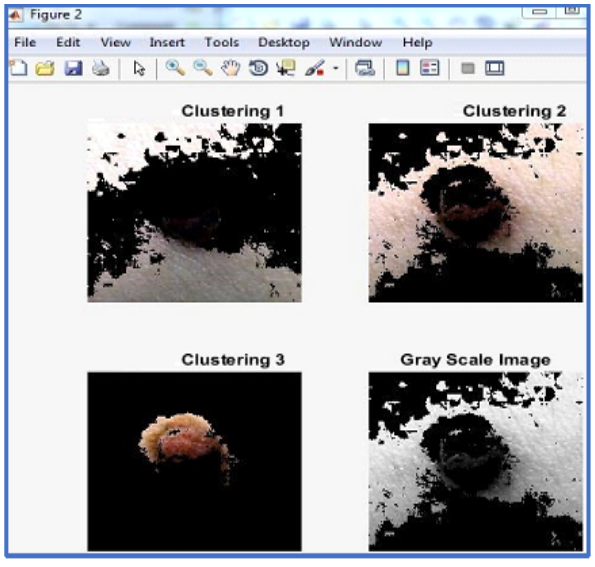
machines (SVM) may be used to accurately identify skin abnormalities using MATLAB. In this study, 812 colored skin images from the Department of Dermatology at Khulna Medical College were used, which included 24 different types of skin diseases, including scabies, psoriasis, seborrheic dermatitis, tinea corporis, chronic eczema, vitiligo and leprosy; these included images of 24 different types of skin diseases. In the dataset, you'll find examples of images of sick skin like the ones you see above. The input/query image is depicted using a distinct color channel (RGB). The histogram of a grayscale image can also be computed. With this technique, we were able to identify and analyze skin diseases that have not previously been studied by scientists, such as basal cell carcinoma and pemphigus. Figure 1 shows the final result of the color alteration of the input image.



(a)



(b)



(c)

Figure 3: Infected skin image segmentation using proposed TSA clustering

The contrast-enhanced method and the Median filter are used to smooth out the noise in the infected image that was provided as input. The original query image is used to build the HIS file. The sick zone, on the other hand, is clearly marked in Figure 3. At the end of the process, hybrid TSA clustering is utilized to label and segment the image. This technique, which employs a

combination of the clustering method and the TSA clustering algorithm, is capable of segmenting the input contaminated skin image and providing the most practicable solution for it. The hybrid TSA approach makes use of three clusters to achieve its results. Pixels from both the diseased and healthy skin areas are detected at this stage, which is accomplished through the use of a sterile barrier. As illustrated in Figure 3, a segmented output of the TSA algorithm has been generated.

For images, TSA training is done during the training phase so that high detection accuracy is achieved. They have been approved into the testing phase. After the segmented picture features have been calculated, the data is compared to the training data. On the other hand, this classifier is able to tell the difference between different types of diseases based on photographs of diseased tissue.

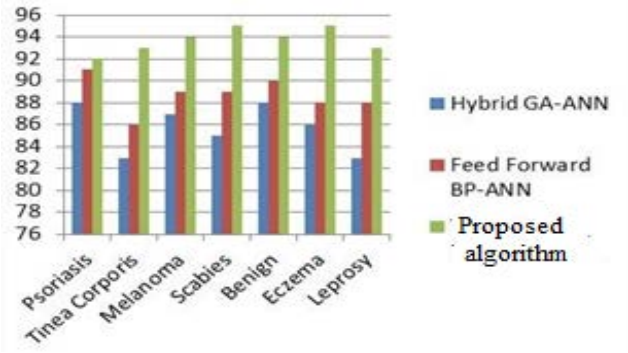


Figure 4: Comparison of Accuracy for Hybrid GA-ANN, Feed Forward BP- ANN and our proposed TSA with RFAT algorithm.

Hybrid GA-ANN and Feed Forward BP-ANN are proposed in two studies in this field for identifying various patterns of skin disorders. Their technique has a detection rate of 88% and 90%, respectively. An inadequate study is being conducted with only 5-9 different types of skin diseases. When it comes to detecting various skin disorders with accuracy levels of up to 95%, our suggested system of RFAT algorithm with TSA classifier has achieved the best results to date. Figure 4 depicts the comparison of accuracy.

7 Conclusion

This section discusses the Regression with the Fuzzy c-mean Adversarial Tunicate algorithm (RFAT) approach, which is based on Oppositional Tunicate Fuzzy C-mean regression to detect cloud intrusions. The performance has been examined using a variety of datasets, including synthetic, NIB, IBRL, and NSL-KDD. The suggested work has been compared to ML-IDDS, En-ABC, Lr-HID and ANN among other state-of-the-art efforts. As an added bonus, the maximum TPR and Low FPR values were computed using a performance assessment that took into account 10 different attributes and took into account assaults like DoS, R2L, Probe, U2R, and normal. Additionally, twenty benchmark functions and the problems they raise were used to determine the proposed work's efficiency. The convergence rate for the suggested and state-of-the-art works such as GWO, En-ABC and TSA was also assessed to analyse the performance and it was found that the proposed work has better convergence results than some other methods. Moreover, the false-positive rate has been estimated and inference that the proposed work has the lowest FPR. Furthermore, the

performance metrics such as accuracy, TPR, and FPR were analyzed for the proposed as well for the existing works and concluded that the proposed work has better outcomes.

8 Availability of Data and Material

Data can be made available by contacting the corresponding author.

9 References

Yoshida H, Dachman AH. Computer-aided diagnosis for CT colonography. In Seminars in Ultrasound, CT and MRI. 2004;25(5):419-31. DOI: 10.1053/j.sult.2004.07.002

Doi, K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. Comput. Med. Imaging Graph. 2007;31:198–211.

Trabelsi O, Tlig L, Sayadi M, & Fnaiech F. Skin disease analysis and tracking based on image segmentation. 2013 International Conference on Electrical Engineering and Software Applications, Hammamet, 1–7

Keke S, Peng Z, & Guohui L. Study on skin color image segmentation used by fuzzy-c-means arithmetic. In 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, Yantai. 612–5

Lu J, Manton JH, Kazmierczak E, & Sinclair R. Erythema detection in digital skin images. In 2010 IEEE International Conference on Image Processing, Hong Kong, 2545-8.

Sumithra R, Suhil M, & Guru DS. Segmentation and classification of skin lesions for disease diagnosis. Proced. Comput. Sci. 2015;45,76–85.

Maglogiannis I, Zafiropoulos E, & Kyranoudis C. Intelligent segmentation and classification of pigmented skin lesions in dermatological images in Advances in Artificial Intelligence. SETN 2006. In Lecture Notes in Computer Science. (Antoniou, G. et al. Eds). 2006;3955:214–23.

Srivastava D, Singh R, Singh V. An intelligent gray wolf optimizer: a nature inspired technique in intrusion detection system (IDS). Journal of Advancement in Robotics. 2019;6(1):18-24.

Garg S, Kaur K, Batra S, Aujla GS, Morgan G, Kumar N, Zomaya AY, Ranjan R. En-ABC: An ensemble artificial bee colony based anomaly detection scheme for cloud environment. Journal of Parallel and Distributed Computing. 2020;135:219-33.

Awan SM, Khan ZA, Aslam M, Mahmood W, Ahsan A. Application of NARX based FFNN, SVR and ANN Fitting models for long term industrial load forecasting and their comparison. In 2012 IEEE International Symposium on Industrial Electronics 2012 (pp. 803-7). IEEE.

Besharati E, Naderan M, Namjoo E. LR-HIDS: logistic regression host-based intrusion detection system for cloud environments. Journal of Ambient Intelligence & Humanized Computing. 2019;10(9):3669-92.

Garg S, Kaur K, Batra S, Aujla GS, Morgan G, Kumar N, Zomaya AY, Ranjan R. En-ABC: An ensemble artificial bee colony based anomaly detection scheme for cloud environment. Journal of Parallel and Distributed Computing. 2020;135:219-33.

Ravinder Reddy Baireddy got his M.Tech From JNTUH. He is a Ph.D student at Vishveshwarayya Technological University. He is an Assistant Professor at Sreenidhi Institute of Science and Technology, Hyderabad. His research involves Data Mining - Dermatologic Diseases Analysis using Swarm Optimization Algorithm.

Dr. R. Nagaraja is a former Professor & PG Coordinator in Bangalore institute of technology in ISE Dept. His research is in the area of Computer Sciences.

Dr. G. Sudha is an Associate Professor & Head of the EEE Department, Bangolore Institute of Technology. Her research involves Power Systems.