An efficient gravitational search decision forest approach for fingerprint recognition

Mahesh Kumar*, Devender Kumar

Dept. of Computer Science and Engineering,
Baba Mastnath University, Asthal Bohar, Sector-29, Rohtak, India
**Corresponding author: malkanimahesh@gmail.com*

Abstract

Fingerprint based human identification is one of the authentic biometric recognition systems due to the permanence and uniqueness of the finger impressions. There is the extensive usage of fingerprint recognition in personalized electronic devices, security systems, banking, forensic labs, and especially in law enforcement agencies. Although the existing systems can recognize fingerprints, they lack in case of poor quality and latent fingerprints. The latent fingerprints are captured by law enforcement agencies during the crime scene to find the criminal. Consequently, it is essential to develop a novel system that can efficiently recognize both complete and latent fingerprints. The current work proposes an efficient Gravitational Search Decision Forest (GSDF) method, which is a combination of the gravitational search algorithm (GSA) and the random forest (RF) method. In the proposed GSDF approach, the mass agent of GSA determines the solution by constructing decision trees in accordance with the random forest hypothesis. The recognition of the fingerprints is accomplished by mass agents in the form of a final generated decision forest from the set of hypothesis space as the mass agents can create multiple hypotheses using random proportional rules. The experiments for fingerprint recognition are conducted for both the latent fingerprints (NIST SD27 dataset) and the complete fingerprints (FVC2004 dataset). The effectiveness of the proposed GSDF approach is analyzed by evaluating the results with machine learning classifiers (random forest, decision tree, back propagation neural networks, and k-nearest neighbor) as well. The comparative analysis of the proposed approach and incorporated machine learning classifiers indicates the outperformed performance of the proposed approach.

Keywords: Back propagation neural networks; decision tree; fingerprint recognition; gravitational search algorithm; k-nearest neighbor; latent fingerprints; machine learning; random forest

1. Introduction

There are numerous biometric systems for human identification, including iris recognition, face recognition, fingerprint recognition, etc. (Nadeem *et al.*, 2022). Among these methods, fingerprint recognition is the most widely adapted method in practice. The concept of fingerprint recognition can be represented in two aspects: verification and identification (Maltoni *et al.*, 2009). The verification aspect is the 1:1 comparison of a human's fingerprints with previously stored data. The identification aspect is the 1:N comparison to determine the identity of a human by comparing the unknown fingerprints with the available overall fingerprint databases. The verification aspect is used for complete fingerprint based biometric systems, and the identification aspect is used by law enforcement agencies to identify the suspect on the basis of acquired latent or complete fingerprints. The current work focuses on both aspects by experimenting with both complete and latent fingerprints. An illustration of latent and complete fingerprints is depicted in Figure 1. Latent fingerprints are poor quality distorted finger impressions, whereas complete fingerprints can be plain or rolled impressions. Plain finger impressions are made by pressing

Fig. 1. Types of Fingerprints.

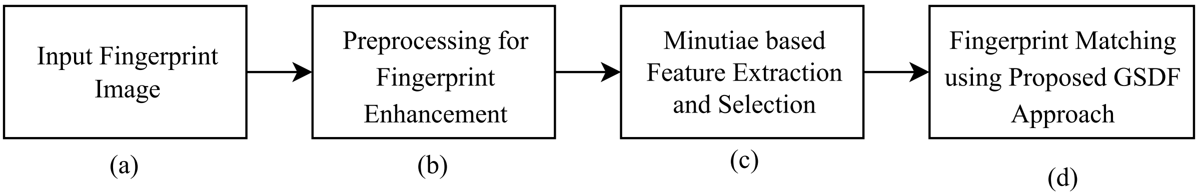


Fig. 2. Process of Fingerprint Recognition (a) Fingerprint Consideration, (b) Preprocessing, (c) Feature Extraction & Selection, and (d) Fingerprint Matching.

the finger on a surface, whereas rolled finger impressions are made by rolling the finger from one side of the fingernail to the other.

The process of fingerprint recognition is discussed by different modules: fingerprint consideration; preprocessing; feature extraction & selection; and fingerprint matching. A brief overview of the fingerprint recognition process is described in Figure 2. For the incorporated fingerprints, the preprocessing module enhances the poor quality and latent fingerprints by using the ridge dictionary and Gabor filter. Further, the minutiae-based features are extracted by using the crossing number concept. In the feature selection phase, the spurious minutiae are removed prior to beginning the fingerprint matching. The final module of fingerprint matching is performed using the proposed GSDF approach, which recognizes the fingerprints by constructing the decision forest with the help of mass agents. The amalgamation of the machine learning based RF algorithm with the GSA algorithm is owing to the stability of the GSA method, which is theoretically modeled using Newton's laws. In addition, the GSA's effective use in several fields of bioinformatics, digital image processing, robotics, and optimization (Kumar *et al.*, 2020) has prompted its use in the present work of fingerprint recognition. The main contributions of the work are summarized as follows.

- The proposal of an efficient GSDF approach by amalgamating GSA and RF algorithms for fingerprint recognition.
- The autonomous enhancement of latent and poor quality fingerprints using a combination of ridge dictionary and Gabor filter.
- The incorporation of a feature selection module to remove spurious minutiae extracted during the minutiae extraction phase. The removal of spurious minutiae enhances recognition accuracy.
- The testing of the proposed approach for both the latent fingerprints (NIST SD27 dataset) and the complete fingerprints (FVC2004 dataset).

The organization of the remaining portions of the paper is as follows: Section 2 presents the state-of-the-art work related to fingerprint recognition. Section 3 describes the preprocessing of the input fingerprint images. Section 4 depicts the minutiae based feature extraction and selection module for fingerprint matching. Section 5 explains the proposed GSDF approach, which is utilized for fingerprint recognition. Section 6 evaluates the results for the experiments on the FVC2004 and NIST SD27 datasets. Also, the comparative analysis of the proposed approach with incorporated machine learning classifiers is conducted in Section 6. Finally, Section 7 illustrates the conclusion of the work along with future directions.

2. Related work

Fingerprint recognition systems should be autonomous and reliable. Inaccurate information might lead to mishaps, especially in the case of latent fingerprints acquired from the crime scene. In 2004, the FBI erroneously detained a man from Oregon in connection with the explosion investigation. This raised problems regarding the fundamentals of forensic science and technology (Newman, 2007). This incident has increased the focus of researchers on the development of efficient fingerprint recognition technologies. Here, the recent studies in the field of fingerprint recognition are discussed.

Guo *et al.* (2014) adapted the decision tree rule-based approach for fingerprint classification. The authors also incorporated the methods of balance arm flow and center to data flow for the recognition of indistinguishable fingerprints. Hsieh & Hu (2014) hybridized the support vector machine (SVM) with particle swarm optimization (PSO) for the classification of fingerprints. The hybridized approach was served with multi-objective optimization to handle the penalty errors of the SVM algorithm. Babatunde (2015) proposed minutiae-based matching for the fingerprints from the different data sources. The spatial and Euclidean relations among the minutiae were evaluated, and pattern matching was conducted from the singular core points. Murugan & Rose (2017) used the back propagation neural network for the recognition of plain and rolled fingerprint images. Lee *et al.* (2017) worked on the recognition of partial fingerprints using ridge shape features (RSF) and minutiae information. The authors designed this algorithm to improve the recognition of fingerprints on small scanning devices such as smart phones.

Cao & Jain (2018) focused on latent fingerprint matching using the convolutional neural network. The feature attributes of minutiae information and texture templates were adapted for the fingerprint feature representation. Castillo-Rosado & Hernández-Palancar (2019) used the distinctive ridge point method for latent fingerprint matching. Wong & Lai (2020) adapted the orientation field information along with the multi-tasking convolutional neural network for the restoration of corrupted fingerprints. Kumar & Garg (2020) introduced the hybrid approach of particle swarm optimization and cuckoo search for latent fingerprint recognition. Jindal & Singla (2021) used an ant colony optimization algorithm for matching the minutiae of latent fingerprints with original fingerprints. Deshpande *et al.* (2021) presented a ratio to minutiae triangles based method which is a rotation and scale invariant approach. The presented method was used for the identification of latent fingerprints. Pradeep & Ravi (2022) incorporated the artificial neural network (ANN) for fingerprint classification after extracting the features using Gabor filter. Singla *et al.* (2022) hybridized the features of pores and minutiae points for the identification of latent fingerprints. Existing studies indicate the usability of different techniques for latent and complete fingerprint recognition systems. This work addresses the following research gaps in some existing studies.

- The focus of the researchers is observed either on the complete fingerprints with some noise value or latent fingerprints with good quality images. There is a need to develop a system that can handle the complete as well as latent fingerprints of low quality images.
- The manual analysis of complex latent fingerprint structures is also challenging for matching with complete fingerprints. The present work autonomously performs the module.
- The recognition accuracy of the existing fingerprint recognition systems should be improved, especially the latent fingerprint recognition, as false values can lead to punishment for any benign person.
- During fingerprint extraction, there may be false minutiae extracted along with the actual minutiae. The removal of spurious minutiae information should also be incorporated as the post processing step to reduce the false positive and false negative rates. The present work also addresses this concern as the feature selection module.

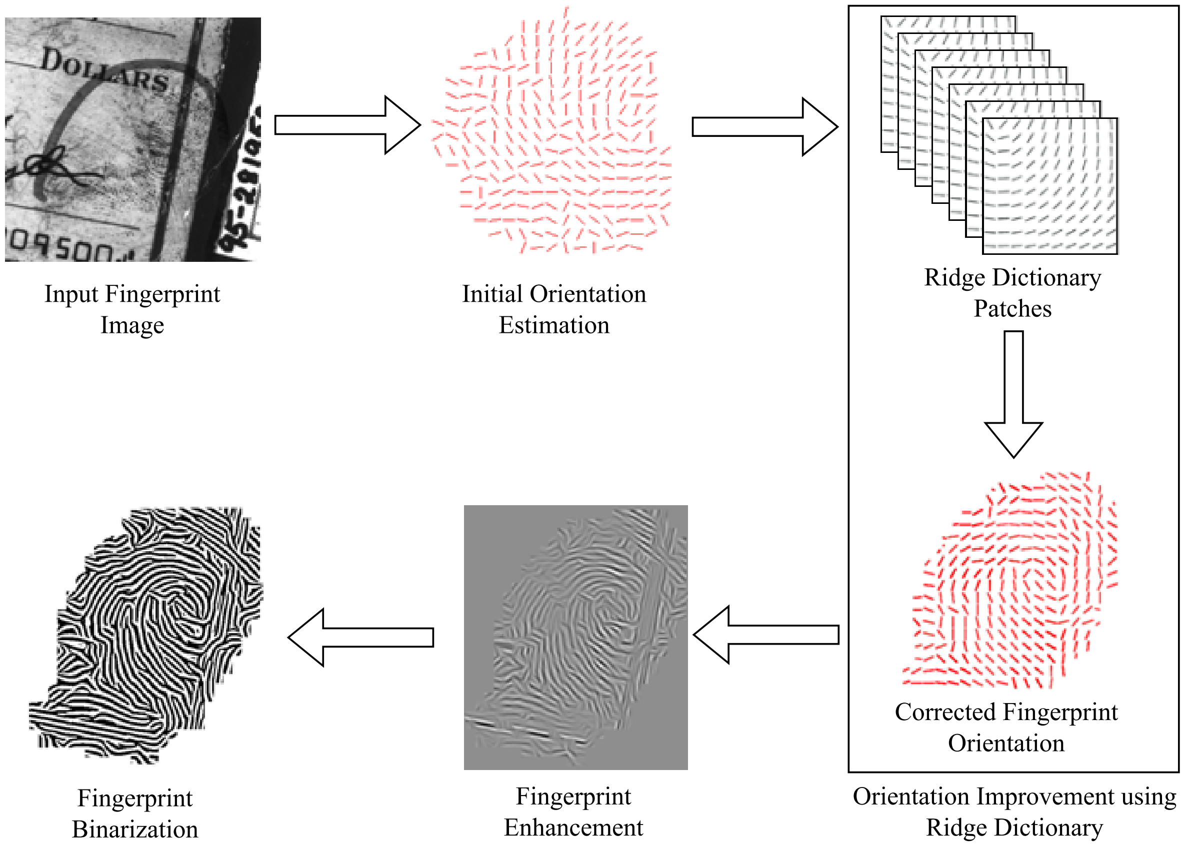


Fig. 3. Preprocessing of Fingerprints.

3. Fingerprint preprocessing

Fingerprint preprocessing is the essential module of the fingerprint recognition process as poor quality fingerprints cannot extract the minutiae features efficiently. Fingerprint preprocessing is composed of four essential steps: fingerprint orientation estimation; orientation improvement using a ridge dictionary; enhancement; and binarization. These steps are also depicted in Figure 3 by considering a latent fingerprint image from the NIST SD27 dataset.

3.1 Fingerprint orientation estimation

Initially, the fingerprint images are segmented and normalized to estimate the orientation of the fingerprints. Segmentation separates the foreground region from the background while preserving the fingerprint ridges and other features. The region of interest (finger impression) from the image is extracted using the variance method. For this process, the image $I(i, j)$ is splitted into 16×16 blocks and the variance $V(I)$ is evaluated for each block. The blocks with a variance value greater than the threshold are retained because the background regions possess a lower threshold value. The obtained finger impression is normalized to reduce the variations in the grey-level of fingerprints while retaining the valley and ridge information unaffected. The normalization $N(i, j)$ of the image at pixel-level is conducted by considering the desired mean and variance values of M_0 and V_0, respectively. The normalized image is processed for orientation estimation using the gradient vectors, which determine local orientation towards the ridge direction flow (Jindal & Singla, 2021).

The orientation image illustrates the invariant coordinates of the fingerprints and analyzes the local ridge information. For the gradient vector method, the normalized image $N(i, j)$ is divided into blocks of size 16×16. At each pixel (i, j) of each block, orientation $O(i, j)$ is estimated with least square estimation using Equation (1).

$$O(i, j) = \frac{1}{2}\tan^{-1}\left(\frac{G_y(i, j)}{G_x(i, j)}\right) \tag{1}$$

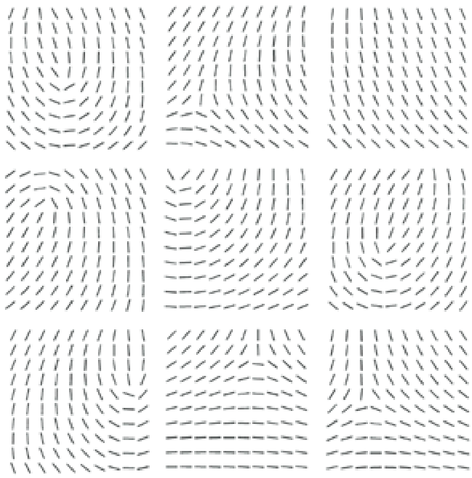


Fig. 4. Samples of Orientation Patches from the NIST SD4 dataset for Ridge Dictionary Construction.

Where, the gradient vector $G_x(i,j)$ and $G_y(i,j)$ are evaluated using the Sobel operator (Hong *et al.*, 1998) for the gradients $\partial_x(i,j)$ and $\partial_y(i,j)$ with respect to the x-axis and y-axis, respectively. The calculated orientation values are kept as matrices.

Due to the low quality of the input latent fingerprint, as seen in Figure 3, the estimated orientation field is noisy. Consequently, orienting is enhanced with the use of ridge dictionary. Further, the ridge dictionary is constructed and the orientation field is smoothed.

3.2 Orientation improvement using ridge dictionary

The ridge dictionary is constructed from the NIST SD4 dataset, which is composed of high-quality rolled fingerprints. The orientation patches, including ridge information, are retrieved from the fingerprints of this dataset with a block size of 16×16 pixels. Each orientation patch consists of 10×10 orientation elements. Figure 4 shows some of the high-quality orientation patches that were taken from the NIST SD4 dataset (Cao & Jain, 2015).

In the constructed ridge dictionary, only the unique patches with a quality index greater than the threshold are included, with no recurrence of ridge patterns. With the addition of the ridge dictionary, fingerprint orientation gets corrected to a great extent. Further, the fingerprint image with corrected ridge orientation is smoothed using a low-pass filter (Jain *et al.*, 2000) in which the image is initially converted to a continuous vector field as depicted by Equations (2)-(3).

$$\Phi_x(i,j) = \cos(2\theta(i,j)) \tag{2}$$

$$\Phi_y(i,j) = \sin(2\theta(i,j)) \tag{3}$$

Where, Φ_x and Φ_y are the vector field components with respect to the x and y axes respectively. As per the low-pass filter, the resulting vector field is determined in terms of Φ'_x and Φ'_y using Equations (4)-(5).

$$\Phi'_x(i,j) = \sum_{u=-w_\Phi/2}^{w_\Phi/2} \sum_{v=-w_\Phi/2}^{w_\Phi/2} W(u,v)\Phi_x(i-uw, j-vw) \tag{4}$$

$$\Phi'_y(i,j) = \sum_{u=-w_\Phi/2}^{w_\Phi/2} \sum_{v=-w_\Phi/2}^{w_\Phi/2} W(u,v)\Phi_y(i-uw, j-vw) \tag{5}$$

Where, $W(u,v)$ is the low-pass filter with a filter size of $w_\Phi \times w_\Phi$. Further, the final ridge orientation O' is estimated using Equation (6).

$$O'(i,j) = \frac{1}{2}\tan^{-1}\frac{\Phi'_y(i,j)}{\Phi'_x(i,j)} \tag{6}$$

The corrected orientation image with the help of the ridge dictionary is illustrated in Figure 3.

3.3 Fingerprint enhancement

Enhancement is conducted to remove the undesired noise and preserve the corrected ridge and orientation information. The attributes of the Gabor filter, such as orientation-selective and frequency-selective, can efficiently remove the noise by preserving the ridge structure and orientation information. Moreover, it is efficient in both the frequency and spatial domains. This makes the Gabor filter a perfect fit for the enhancement process. The formulation of the Gabor filter (Hong *et al.*, 1998) in the spatial domain is described by Equations (7)-(9).

$$H(x,y;f,\phi) = \exp\left\{-\frac{1}{2}\left[\frac{x_\phi^2}{\delta_x^2}+\frac{y_\phi^2}{\delta_y^2}\right]\right\}\cos(2\pi f x_\phi) \tag{7}$$

$$x_\phi = x\cos\phi + y\sin\phi \tag{8}$$

$$y_\phi = -x\sin\phi + y\cos\phi \tag{9}$$

Where, f is the filter frequency, and ϕ is the orientation of the Gabor filter. δ_x and δ_y are the standard deviations with respect to axes x and y.

3.4 Binarization

The binarization process transforms the grey-level filtered image into a binary image. Here, the local adaptive binarization approach is adapted for transformation in which the mean intensity value is evaluated as a threshold value. The final image is obtained by assigning the value of 1 to the pixels whose values are higher than the threshold, and the assigning value 0 to the rest of the pixels.

4. Feature extraction and selection

For fingerprint recognition, minutiae-based features are extracted, which are specific to normal pixel, ridge bifurcation, and ridge endings. Here, the crossing number method is utilized to evaluate the minutiae-based features. It determines the feature types by analyzing the surrounding pixels of a pixel P within the 3×3 pixel window. The finding of crossing numbers 1, 2, 3, or greater than 3 reveals the minutiae features of ridge ending, a normal ridge pixel, and ridge bifurcation, respectively. Figure 5 depicts the assessment of features using the crossing number approach. These minutiae feature types are extracted for fingerprint matching.

Prior to considering the extracted minutiae features for fingerprint matching, these features are filtered to exclude any irrelevant minutiae. The feature selection procedure eliminates spurious minutiae such as dots, ladders, lakes, triangles, breaks, etc. Here, the ridge dictionary is utilized to identify the spurious minutiae. The feature selection is required since spurious minutiae might lead to erroneous fingerprint matching. The selected feature set is stored for fingerprint matching. A sample of minutiae features after the feature selection is illustrated in Figure 6.

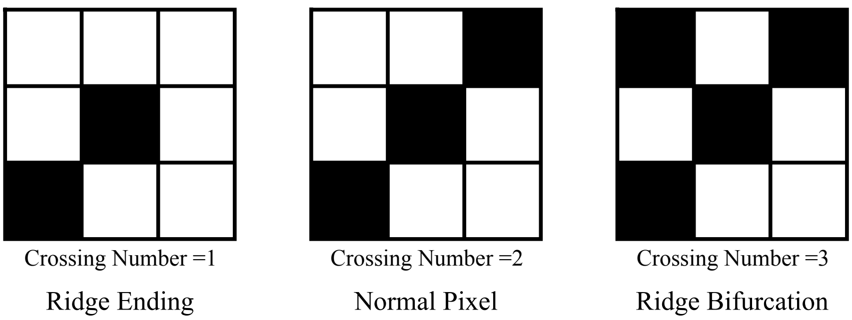


Fig. 5. Minutiae Feature Types using the Crossing Number Method.

5. Fingerprint matching using proposed GSDF approach

The fingerprint matching is conducted using the proposed GSDF approach, which is an amalgamation of GSA and RF algorithms. The GSA algorithm is a physics-inspired meta-heuristic algorithm that follows Newton's laws of gravity and motion for optimization (Jindal *et al.*, 2022). The RF algorithm is an ensemble of decision trees constructed with randomly selected characteristics (Manpreet & Chhabra, 2022). In the proposed GSDF approach, GSA's mass agents construct decision trees by following the hypothesis of a random forest algorithm in which random solutions are generated based on splitting rules and thresholds. The mass agents also determine the new random sub-space to handle the increasing decision trees and keep the tradeoff between exploration and exploitation. This amalgamation process conducts the fingerprint matching with better accuracy compared to the decision tree alone (Kozak, 2019). The process of GSDF approach initiated by considering N number of mass agents, with the initial position of i^{th} agent as X_i, the initial gravitational constant G_0, and a decision table with decision attributes d_a. The initial force acting (Rashedi *et al.*, 2009) on the agent i by the agent j in d-dimensions is determined by Equation (10). The mass agents describe the pixels of the fingerprint image and construct an overall decision forest to determine which fingerprint in the fingerprint database matches the input fingerprint.

$$F_{ij}^d(t) = G(t)\frac{M_{pi}(t) \times M_{aj}(t)}{R_{ij}(t) + \varepsilon}\left(x_j^d(t) - x_i^d(t)\right) \tag{10}$$

Where, $G(t)$, M_{pi}, and M_{aj} are the gravitational constant, passive mass for agent i, and active mass for agent j respectively. The term ε is constant and R_{ij} is the Euclidean distance between mass agents. The addition of stochastic attributes to the GSDF upgrades the Force on mass agents as depicted by Equation (11).

$$F_i^d(t) = \sum_{j=1, j\neq i}^{N} rand_j F_{ij}^d(t) \tag{11}$$

Where, $rand_j$ is a random number in the range [0, 1].

For the movement of the mass agents in nodes to construct the decision trees, the acceleration value is also evaluated by following Newton's law of motion. The formula to evaluate the acceleration $a_i^d(t)$ is depicted by Equation (12).

$$a_i^d(t) = \frac{F_i^d(t)}{M_{ii}(t)} \tag{12}$$

Where, $M_{ii}(t)$ is the inertial mass of i^{th} agent.

As the inertial and gravitational masses are computed using the fitness function, which states that a higher mass value indicates a superior agent. For a better solution space with heavy masses, the inertial and gravitational masses are equalized. This updates the masses as described in Equations (13)-(14).

$$M_{ii} = M_{pi} = M_{ai} = M_i \tag{13}$$

$$M_i(t) = \frac{m_i(t)}{\sum_{j=1}^{N} m_j(t)} \tag{14}$$

Where, the value of $m_i(t)$ is evaluated (Equation (15)) by considering the best ($best(t)$) and worst ($worst(t)$) values for mass agents.

$$m_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \tag{15}$$

Where, $fit_i(t)$ is the fitness function.

Each mass agent constructs the decision tree by adapting the random attributes of the RF algorithm. Further, the generated multiple decision trees are ensembled and a final decision is made for the fingerprint matching. There is a test on the attributes of each node of the decision tree as depicted by Equation (16).

$$test : O \to R_{test} \tag{16}$$

Where, the set of objects is defined by O and the possible tests are annotated with $R_{test} = \{r_1, r_2, \ldots, r_z\}$. Further, the applicability of the test for the attributes $a : O \rightarrow A$ is described by Equation (17).

$$test : A \rightarrow R_{test} \tag{17}$$

Here, the possible sub-trees $(T_1, T_2, \ldots, T_z)$ can be constructed by each node which tests $(r_1, r_2, \ldots, r_z)$ for the consideration of the assumption that T_i sub-trees are created by test r_i. This derives the hypothesis $h(x)$, as shown in Equation (18).

$$h(x) = \begin{cases} h_1(x), test(x) = r_1 \\ h_2(x), test(x) = r_2 \\ \vdots \\ h_z(x), test(x) = r_z \end{cases} \tag{18}$$

For the n number of nodes, the size of constructed decision tree is evaluated using Equation (19).

$$s(T) = \frac{1}{n} \tag{19}$$

In the GSDF approach, the heuristic function η_{A_i,V_j} for the attributes A_i and values V_j is calculated by following the Twoing splitting criteria to attain the best split of the tree. It also retains the maximum homogeneity of the nodes in the tree (Vives *et al.*, 2021). The formula for the evaluation of η_{A_i,V_j} is described by Equation (20).

$$\eta_{A_i,V_j} = \frac{P_l P_r}{4} \left[\sum_{d=1}^{D} \left| p\left(d|node_{l(A_i,V_j)}\right) - p\left(d|node_{r(A_i,V_j)}\right) \right| \right]^2 \tag{20}$$

Where, D is the maximum number of possible decision classes, P_l and P_r are the probabilities for the left and right nodes. $p\left(d|node_{l(A_i,V_j)}\right)$ and $p\left(d|node_{r(A_i,V_j)}\right)$ are the conditional probabilities for the left and right nodes, respectively.

The movement of the mass agents from one node to another makes it necessary to determine the updated position and velocity of agents. The changes in position and velocity values as per the GSA algorithm are determined by Equations (21)-(22).

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \tag{21}$$

$$v_i^d(t+1) = rand_i \times v_i^d(t) + a_i^d(t) \tag{22}$$

The process of construction of the decision tree continues, and iterations are also increments. In the later iterations, the mass agents can be trapped in the local optimum due to the heaviness of the masses with the increasing iterations. This situation is handled by introducing the function of $Kbest$ which is a function of time. The $Kbest$ agents also possess the highest mass value, the best fitness, and it decreases linearly with time. At the end, there will be the applicability of force by one agent to others, and the change in force is described by Equation (23).

$$F_i^d\left(t\right) = \sum_{j \epsilon Kbest, j \neq i} rand_j F_{ij}^d\left(t\right) \tag{23}$$

To determine the final match for the fingerprints, the outcome of each decision tree is analyzed which will be further ensembled to determine the outcome of the decision forest by following the attributes of the RF algorithm. The fingerprint classification and recognition outcome by each decision tree $(T(S))$ with training sample (S) is determined by Equations (24)-(25).

$$\epsilon\left(T\left(S\right), Dst\right) = \sum_{(x,y) \epsilon U} Dst\left(x, y\right).L\left(y, T\left(S\right)\left(x\right)\right) \tag{24}$$

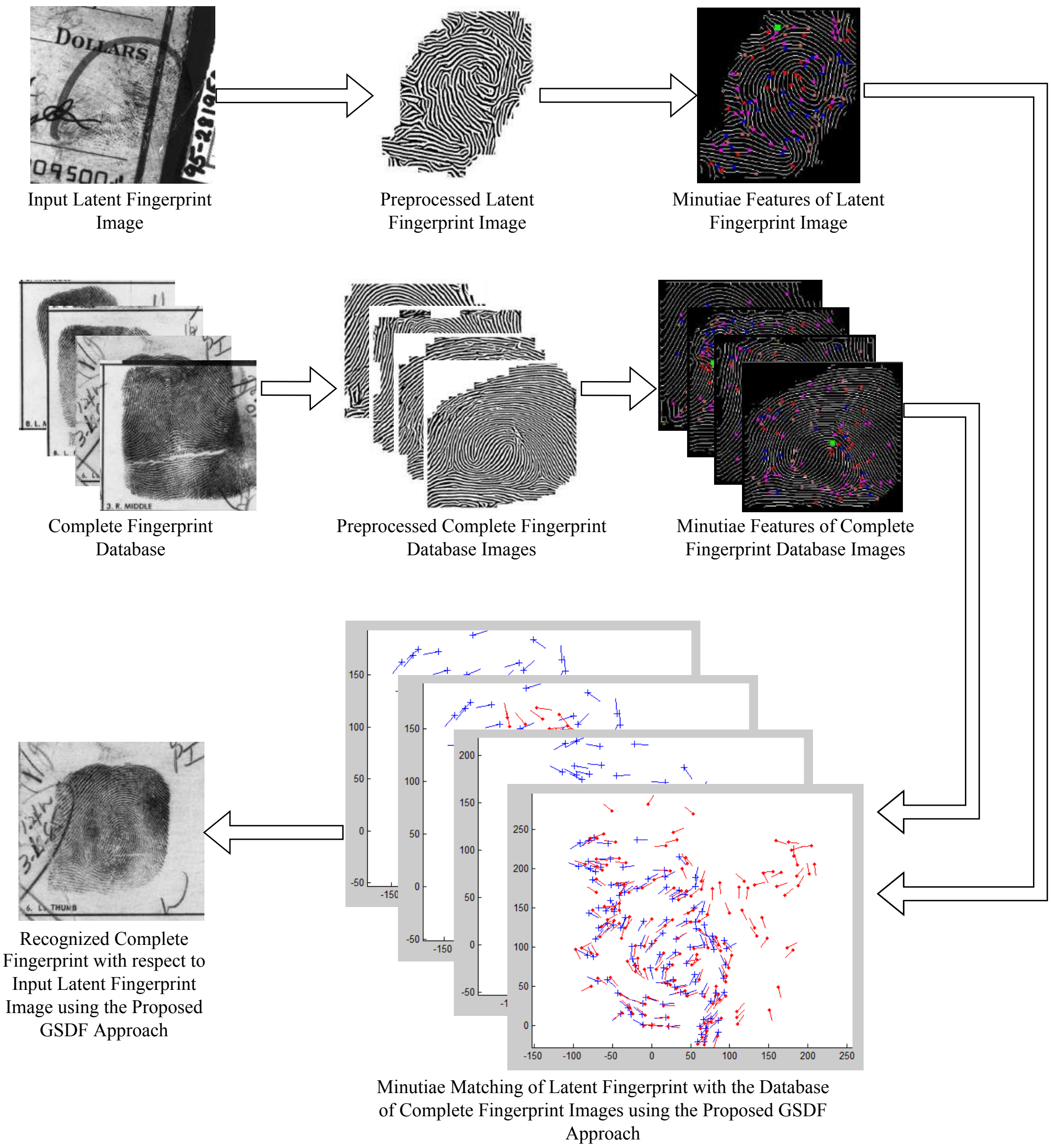


Fig. 6. Fingerprint Recognition using the Proposed GSDF Approach.

$$L\left(y, T\left(S\right)\left(x\right)\right) = \begin{cases} 1, & \text{if } y \neq T\left(S\right)\left(x\right) \\ 0, & \text{if } y = T\left(S\right)\left(x\right) \end{cases} \tag{25}$$

Where, the possible values of attributes are denoted by U and the distribution is denoted by Dst.

The overall results of the fingerprint classification are evaluated as a decision forest by combining the outcomes of the decision trees with the help of voting criteria. The availability of diversity in attributes of the GSDF approach makes the agents to choose different nodes for the construction of decision trees with different combinations, hence the ensemble decision forest. The final solution set is determined by the completion of maximum iterations and the evaluation of the solution by all the mass agents. The pseudo-code of the proposed GSDF approach for fingerprint recognition is described by Algorithm 1. The pictorial representation of the fingerprint recognition using the proposed GSDF approach is described in Figure 6.

Algorithm 1: Pseudo-Code of the Proposed GSDF Approach for Fingerprint Recognition

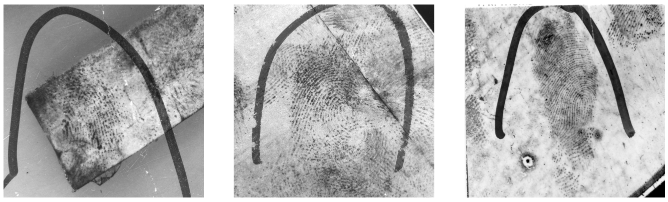
```
Initialize the parameters of the GSA and RF algorithms.
decision_forest=null;
iteration=1;
while iteration ≤ iteration_max do
    for (j = 1 to number_of_decision_trees) do
        best_decision_tree=null;
        fingerprint_classifier=choose_objects // Consider mass agents for the pixels of the
          fingerprint image data with equal probability.
        for (N = 1 to number_of_mass_agents) do
            Construct decision trees by considering subset of attributes at each node using
              attributes of GSDF approach.
            new_decision_tree=decision_tree_construction_using_GSDF_attributes.
            if (new_decision_tree_quality) > (best_decision_tree_quality) then
                best_decision_tree = new_decision_tree;
            end
        end
        Update position and velocity of mass agents.
        decision_forest.add (best_decision_tree);
    end
    iteration=iteration+1;
end
Outcome=decision_forest // with final classification and recognition of fingerprints.
```

6. Results and discussion

The fingerprint recognition results are evaluated for the latent fingerprint dataset of NIST SD27 and the complete fingerprint dataset of FVC2004. In the NIST SD27 dataset (Garris & McCabe, 2000), a total of 258 fingerprint images along with their rolled fingerprints are available. These latent fingerprints are available in three categories: ugly, bad, and good, with respective images of 85, 85, and 88. Further, the FVC2002 dataset is a composition of four sub-datasets of DB1, DB2, DB3, and DB4, collected using various sensors and technologies (Maio *et al.*, 2002). Each sub-dataset consists of 80 fingerprint images. The sample images of the NIST SD27 and FVC2004 datasets are illustrated in Figure 7.

The proposed GSDF approach determines the match of fingerprints by evaluating the similarity score of minutiae features. For the latent fingerprints (NIST SD27 dataset), the minimum threshold value of similarity score is considered to be 75% as the latent fingerprints are incomplete fingerprints. On the other hand, the similarity score is set to be 95% for the complete fingerprints (FVC2004 dataset). The attainment of the mentioned similarity score threshold signifies the accurate match of the fingerprints. The performance results are calculated in terms of precision, recall, f-measure, and recognition rate by using similarity score values. To analyze the effectiveness of the proposed approach, the results are also calculated for the machine learning algorithms of random forest (RF) (Manpreet & Chhabra, 2022), decision tree (DT) (Azad *et al.*, 2022), back propagation neural networks (BPNN) (Kiran *et al.*, 2021), and k-nearest neighbor (KNN) (Manpreet & Chhabra, 2022). Tables 1-3 present the performance evaluation results for the NIST SD27 dataset, and Tables 4-7 describe the performance evaluation results for the FVC2004 dataset.

The results depicted in Tables 1-3 for latent fingerprint (NIST SD27 dataset) recognition indicate that the proposed GSDF approach has matched the latent fingerprints with complete fingerprints efficiently. The proposed approach has attained the recognition rate of 87.06% for the ugly class, 91.76% for the bad class, and 98.86% for the good class of latent fingerprints. The incorporated machine learning algorithms have also matched the latent fingerprint with complete fingerprints, but performance is inferior to the proposed GSDF approach.



Ugly Bad Good

(a) NIST SD27 Dataset

DB1 DB2 DB3 DB4

(b) FVC2004 Dataset

Fig. 7. Sample Images of the (a) NIST SD27 Dataset, and (b) FVC2004 Dataset.

Table 1. Performance Evaluation Results for the Ugly Fingerprint Class of the NIST SD27 Dataset.

Method	Precision (%)	Recall (%)	F-Measure (%)	Recognition Rate (%)
GSDF (Proposed)	72.55	87.06	79.14	87.06
RF	61.11	77.65	68.39	77.65
DT	54.46	71.76	61.92	71.77
BPNN	47.46	65.88	55.17	65.88
KNN	43.90	63.53	51.92	63.53

Table 2. Performance Evaluation Results for the Bad Fingerprint Class of the NIST SD27 Dataset.

Method	Precision (%)	Recall (%)	F-Measure (%)	Recognition Rate (%)
GSDF (Proposed)	76.47	91.76	83.42	91.76
RF	64.81	82.35	72.54	82.35
DT	60.91	78.82	68.72	78.82
BPNN	56.36	72.94	63.59	72.94
KNN	58.93	77.65	67.01	77.65

Table 3. Performance Evaluation Results for the Good Fingerprint Class of the NIST SD27 Dataset.

Method	Precision (%)	Recall (%)	F-Measure (%)	Recognition Rate (%)
GSDF (Proposed)	83.65	98.86	90.63	98.86
RF	76.47	88.64	82.11	88.64
DT	70.48	84.09	76.68	84.09
BPNN	69.23	81.82	75	81.82
KNN	72.12	85.23	78.13	85.23

The results for the complete fingerprint recognition illustrated in Tables 4-7 also indicate that the proposed GSDF approach is more efficient than incorporated machine learning algorithms. For the FVC2004 dataset, the proposed approach has attained a recognition rate of 98.75% for the DB1 class,

Table 4. Performance Evaluation Results for the DB1 Class of the FVC2004 Dataset.

Method	Precision (%)	Recall (%)	F-Measure (%)	Recognition Rate (%)
GSDF (Proposed)	96.34	98.75	97.53	98.75
RF	89.16	92.5	90.80	92.5
DT	83.53	88.75	86.06	88.75
BPNN	82.14	86.25	84.15	86.25
KNN	85.71	90	87.81	90

Table 5. Performance Evaluation Results for the DB2 Class of the FVC2004 Dataset.

Method	Precision (%)	Recall (%)	F-Measure (%)	Recognition Rate (%)
GSDF (Proposed)	96.30	97.5	96.89	97.5
RF	90.12	91.25	90.68	91.25
DT	83.72	90	86.75	90
BPNN	80	85	82.42	85
KNN	83.33	87.5	85.37	87.5

Table 6. Performance Evaluation Results for the DB3 Class of the FVC2004 Dataset.

Method	Precision (%)	Recall (%)	F-Measure (%)	Recognition Rate (%)
GSDF (Proposed)	91.46	93.75	92.59	93.75
RF	83.33	87.5	85.37	87.5
DT	75.58	81.25	78.31	81.25
BPNN	77.91	83.75	80.72	83.75
KNN	72.41	78.75	75.45	78.75

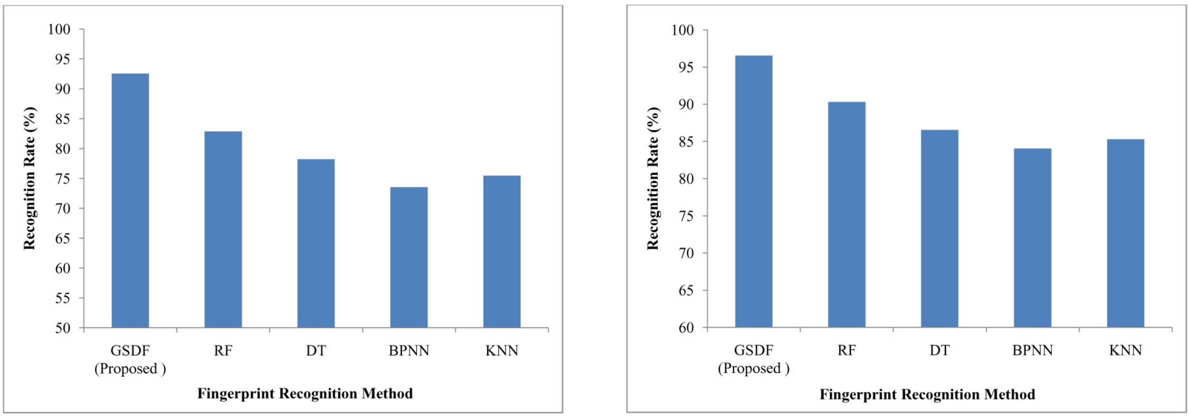
Table 7. Performance Evaluation Results for the DB4 Class of the FVC2004 Dataset.

Method	Precision (%)	Recall (%)	F-Measure (%)	Recognition Rate (%)
GSDF (Proposed)	93.90	96.25	95.06	96.25
RF	85.71	90	87.81	90
DT	80.23	86.25	83.13	86.25
BPNN	74.71	81.25	77.84	81.25
KNN	79.07	85	81.93	85

97.5% for the DB2 class, 93.75% for the DB3 class, and 96.25% for the DB4 class of the dataset which are superior to machine learning algorithms.

Furthermore, the overall comparison of the proposed approach with machine learning algorithms is conducted by incorporating the parameter of recognition rate. For overall comparison, the mean of the results for all the categories of the NIST SD27 and FVC2004 datasets is calculated separately. The comparative analysis is described by Figure 8.

In the overall results, the proposed approach has attained the recognition rate of 92.56% for latent fingerprints and 96.56% for complete fingerprints. For NIST SD27 dataset (Figure 8a), the recognition rate of the proposed GSDF approach is 9.68% better than the RF algorithm, 14.33% better than the DT algorithm, 19.01% better than the BPNN algorithm, and 17.09% better than the KNN algorithm. For FVC2004 dataset (Figure 8b), the recognition rate of the proposed GSDF approach is 6.25% better than the RF algorithm, 10% better than the DT algorithm, 12.5% better than the BPNN algorithm, and 11.25% better than the KNN algorithm. These comparative analysis results clearly indicate that the proposed GSDF approach is efficient compared to incorporated machine learning algorithms for both the latent and complete fingerprints.



(a) **(b)**

Fig. 8. Performance Comparison of the Proposed GSDF Approach with Machine Learning Algorithms for Experiments on (a) NIST SD27 Dataset (b) FVC2004 Dataset.

7. Conclusion

The paper has presented an automated fingerprint recognition system with the proposal of a novel GSDF approach. Initially, the fingerprints from the considered datasets (NIST SD27 and FVC2004) are preprocessed to enhance the poor quality images using a combined ridge dictionary and Gabor filter approach. Further, the minutiae-based features are extracted and spurious minutiae are filtered. The selected features feed into the proposed GSDF approach for fingerprint matching. The proposed approach efficiently determines the match of the fingerprints by constructing the decision trees using mass agents following the hypothesis of random forest. The final fingerprint match is determined by combining the outcomes of all the decision trees. The proposed approach has attained an average recognition rate of 92.56% for latent fingerprints (NIST SD27 dataset) and 96.56% for complete fingerprints (FVC2004 dataset), which are superior to incorporated machine learning algorithms of RF, DT, KNN, and BPNN.

Although the proposed approach has yielded efficient performance results for fingerprint recognition, the recognition rate for ugly latent fingerprints can be further optimized. It will also boost the overall performance of the proposed approach. In the future, we will combine the GSA method with a more effective classifier to improve the performance of ugly quality of the latent fingerprint.

References

Azad, M. & Moshkov, M. (2022). A bi-criteria optimization model for adjusting the decision tree parameters. *Kuwait Journal of Science*, **49**(2), 1-14.

Babatunde, I. G. (2015). Fingerprint matching using minutiae-singular points network. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, **8**(2), 375-388.

Cao, K. & Jain, A. (2015). Latent orientation field estimation via convolutional neural network. In *2015 International Conference on Biometrics (ICB)* (pp. 349-356). Phuket, Thailand: IEEE.

Cao, K. & Jain, A. K. (2018). Automated latent fingerprint recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**(4), 788-800.

Castillo-Rosado, K. & Hernández-Palancar, J. (2019). Latent fingerprint matching using distinctive ridge points. *Informatica*, **30**(3), 431-454.

Deshpande, U. U., Malemath, V. S., Patil, S. M. & Chaugule, S. (2021). Latent fingerprint identification system based on a local combination of minutiae feature points. *SN Computer Science*, **2**(3), 1-17.

Garris, M. D. & McCabe, R. M. (2000). Fingerprint minutiae from latent and matching tenprint images. *National Institute of Standards and Technology*, (1-36).

Guo, J. M., Liu, Y. F., Chang, J. Y. & Lee, J. D. (2014). Fingerprint classification based on decision tree from singular points and orientation field. *Expert Systems with Applications*, **41**(2), 752-764.

Hong, L., Wan, Y. & Jain, A. (1998). Fingerprint image enhancement: algorithm and performance evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**(8), 777-789.

Hsieh, C. T. & Hu, C. S. (2014). Fingerprint recognition by multi-objective optimization PSO hybrid with SVM. *Journal of Applied Research and Technology*, **12**(6), 1014-1024.

Jain, A.K., Prabhakar, S., Hong, L. & Pankanti, S. (2000). Filterbank-based fingerprint matching. *IEEE Transactions on Image Processing*, **9**(5), 846-859.

Jindal, R. & Singla, S. (2021). Ant colony optimisation for latent fingerprint matching. *International Journal of Advanced Intelligence Paradigms*, **19**(2), 161-184.

Jindal, S., Sachdeva, M. & Kushwaha, A. K. S. (2022). Quantum behaved intelligent variant of gravitational search algorithm with deep neural networks for human activity recognition. *Kuwait Journal of Science*, 1-19. DOI: https://doi.org/10.48129/kjs.18531.

Kiran, P., Parameshachari, B. D., Yashwanth, J. & Bharath, K. N. (2021). Offline signature recognition using image processing techniques and back propagation neuron network system. *SN Computer Science*, **2**(3), 1-8.

Kozak, J. (2019). Ant colony decision forest approach. In *Decision Tree and Ensemble Learning Based on Ant Colony Optimization* (pp. 119-134). Cham: Springer.

Kumar, T. & Garg, R. S. (2020). The recognition of latent fingerprints using swarm intelligence based hybrid approach. *International Journal on Emerging Technologies*, **11**(5), 90-97.

Kumar, Y., Verma, S. K. & Sharma, S. (2020). Quantum-inspired binary gravitational search algorithm to recognize the facial expressions. *International Journal of Modern Physics C*, **31**(10), 2050138(1-24).

Lee, W., Cho, S., Choi, H. & Kim, J. (2017). Partial fingerprint matching using minutiae and ridge shape features for small fingerprint scanners. *Expert Systems with Applications*, **87**, 183-198.

Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L. & Jain, A. K. (2002). FVC2002: Second fingerprint verification competition. In *2002 International Conference on Pattern Recognition* (pp. 811-814). Quebec City, QC: IEEE.

Maltoni, D., Maio, D., Jain, A. K. & Prabhakar, S. (2009). Handbook of fingerprint recognition. Springer Science & Business Media, London.

Manpreet & Chhabra J. K. (2022). A hybrid approach based on k-nearest neighbors and decision tree for software fault prediction. *Kuwait Journal of Science*, 1-12. DOI: https://doi.org/10.48129/kjs.18331.

Murugan, A. & Rose, P. A. L. (2017). Fingerprint matching through back propagation neural network. *Indian Journal of Science and Technology*, **10**(29), 1-7.

Nadeem, A., Ashraf, M., Rizwan, K., Qadeer, N., AlZahrani, A., Mehmood, A. & Abbasi, Q. H. (2022). A Novel Integration of Face-Recognition Algorithms with a Soft Voting Scheme for Efficiently Tracking Missing Person in Challenging Large-Gathering Scenarios. *Sensors*, **22**, 1153(1-24).

Newman, D. (2007). The limitations of fingerprint identifications. *Criminal Justice*, **22**, 36.

Pradeep, N. R. & Ravi, J. (2022). An Efficient Machine Learning Approach for Fingerprint Authentication Using Artificial Neural Networks. In *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 852-859). Tirunelveli, India: IEEE.

Rashedi, E., Nezamabadi-Pour, H. & Saryazdi, S. (2009). GSA: a gravitational search algorithm. *Information Sciences*, **179**(13), 2232-2248.

Singla, N., Kaur, M. & Sofat, S. (2022). Hybrid framework for identifying partial latent fingerprints using minutiae points and pores. *Multimedia Tools and Applications*, **81**, 19525–19542.

Vives, L., Tuteja, G. S., Manideep, A. S., Jindal, S., Sidhu, N., Jindal, R. & Bhatt, A. (2021). A novel hybrid approach of gravitational search algorithm and decision tree for twitter spammer detection. *International Journal of Modern Physics C*, 2250060(1-23).

Wong, W. J. & Lai, S. H. (2020). Multi-task CNN for restoring corrupted fingerprint images. *Pattern Recognition*, **101**, 107203(1-11).

Submitted: 24/05/2022
Revised: 16/07/2022
Accepted: 21/07/2022
DOI: 10.48129/kjs.20635