Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
A radiological image analysis framework for early screening of the COVID-19 infection: A computer vision-based approach

Shouvik Chakraborty *, Kalyani Mali

Department of Computer Science and Engineering, University of Kalyani, India

ARTICLE INFO

Article history:
Received 6 September 2020
Received in revised form 20 October 2021
Accepted 21 January 2022
Available online 3 February 2022

Keywords:
COVID-19
Radiological image interpretation
Superpixel
Type 2 fuzzy system
Artificial cell swarm optimization

ABSTRACT

Due to the absence of any specialized drugs, the novel coronavirus disease 2019 or COVID-19 is one of the biggest threats to mankind. Although the RT-PCR test is the gold standard to confirm the presence of this virus, some radiological investigations find some important features from the CT scans of the chest region, which are helpful to identify the suspected COVID-19 patients. This article proposes a novel fuzzy superpixel-based unsupervised clustering approach that can be useful to automatically process the CT scan images without any manual annotation and helpful in the easy interpretation. The proposed approach is based on artificial cell swarm optimization and will be known as the SUFACSO (Superpixel based Fuzzy Artificial Cell Swarm Optimization) and implemented in the Matlab environment. The proposed approach uses a novel superpixel computation method which is helpful to effectively represent the pixel intensity information which is beneficial for the optimization process. Superpixels are further clustered using the proposed fuzzy artificial cell swarm optimization approach. So, a twofold contribution can be observed in this work which is helpful to quickly diagnose the patients in an unsupervised manner so that, the suspected persons can be isolated at an early phase to combat the spread of the COVID-19 virus and it is the major clinical impact of this work. Both qualitative and quantitative experimental results show the effectiveness of the proposed approach and also establish it as an effective computer-aided tool to fight against the COVID-19 virus. Four well-known cluster validity measures Davies–Bouldin, Dunn, Xie–Beni, and $\beta$ index are used to quantify the segmented results and it is observed that the proposed approach not only performs well but also outperforms some of the standard approaches. On average, the proposed approach achieves 1.709792, 1.473037, 1.752433, 1.709912 values of the Xie–Beni index for 3, 5, 7, and 9 clusters respectively and these values are significantly lesser compared to the other state-of-the-art approaches. The general direction of this research is worthwhile pursuing leading, eventually, to a contribution to the community.

1. Introduction

Automated computer-aided systems prove their effectiveness and real-life applicability in various scenarios. Automated systems have a diverse domain of applications and sometimes, these systems are inevitable to perform certain jobs efficiently and in a cost-effective and highly time-bound manner. This domain is evolving day-by-day and continuous effort can be observed from various researchers to enhance this domain. Computer-assisted systems can be categorized in two ways. The first one is the supervised approach in which some properly annotated data are required to perform the classification and interpretation job [1,2]. Therefore, these automated systems are dependent on the ground truth data typically produced by some domain experts. But, it may not be always possible to acquire the properly annotated ground truth data due to the involvement of human experts [3]. Sometimes, some cases are not well-defined or not seen earlier, and therefore, it is very difficult to get some ground truth data for those cases. Unsupervised systems can be helpful in this context because these systems are not dependent on the ground truth data and can automatically explore some patterns from the underlying dataset by utilizing the surrounding knowledge [4–7]. So, the unsupervised approaches are helpful in those situations where a sufficient amount of properly annotated ground truth data are not available. The unsupervised computer-aided systems are widely applied in different domains of research [8,9]. Biomedical image analysis is no exception and exploits the advantages of unsupervised automated systems in various phases. Radiology is one of the important and frequently used parts of the biomedical imaging domain which is serving as an important

https://doi.org/10.1016/j.asoc.2022.108528
1568-4946/© 2022 Elsevier B.V. All rights reserved.
tool for noninvasive diagnostic systems. X-ray, CT Scan, etc. are widely used nowadays, to study the internal functionalities and the present state of the different organs [10–12]. Automated systems are helpful to analyze and diagnose different patients automatically and automated radiological image analysis systems are also helpful in preparing precise and timely reports by reducing the human intervention and also reducing some unintentional human-made errors. Physicians, radiological technicians, and all other concerned domain experts can be significantly benefited from the advancement in the field of computer-aided radiological image analysis systems. Apart from the automated analysis of the radiological images, computer-assisted systems can be helpful in parameter tuning of the image acquisition hardware, image preprocessing, quality control, selecting the appropriate level of radiation, and many more. Therefore, automated systems can act as a helping hand in the decision-making process.

In Table 1 some of the related biomedical image segmentation works of literature are discussed which is helpful in a better understanding of the current trend and status of the same. Apart from these works, some comprehensive studies can be found in [13–17].

Apart from these works, some of the most recent and relevant works can be found in [28–32] that can be referred to, to understand the further advancements of this domain. In this context, it is worth mentioning here that the active contour model is an effective way of image segmentation. There are several variations available of this approach. The traditional active contour approach was proposed in 1988 [33]. A modified version of the traditional active contour approach is proposed in [34] and it is known as geometric active contours. This approach uses gradient information of an image to construct the edge stop function. A region information-based approach is proposed in [35]. This approach is developed by Chan and Vese and this is a parametric representation. Some deep learning approaches are developed that use the loss function of the active contour model as their loss function [36]. Some recent developments like the MAC model [37], SBGFRLS model [38], LSAC model [39], etc. can also be observed in this domain.

The highly infectious coronavirus disease 2019 or COVID-19 creates a worldwide pandemic scenario. Although the mortality rate is not very high, the highly infectious nature of this virus is the main threat to society. Due to the absence of any specialized drug, it is very difficult to restrict the drastic spread of this virus. Apart from using various protective equipment, early detection and isolation can be very effective to combat the spread of this highly infectious virus. In the middle of this pandemic scenario, some vaccines are invented and are being applied to the people and it is ray of hope to fight against this virus. As per the report of the world health organization, 239,437,517 numbers of confirmed cases can be observed in 216 countries and 4,879,235 people are already expired due to this disease as of 15th October 2021, 4:32pm CEST [40]. From these statistics, it is clear that the worldwide mortality rate is approximately 2.0378% which is not a very large value. The major risk factor lies in the highly infectious nature of this virus. Hopefully, 6,495,672,032 vaccine doses have already been administered worldwide which may be helpful in reducing the mortality rate. Many countries are not prepared with the appropriate infrastructures to support COVID-19 infected patients. Moreover, many people from remote areas are not even able to arrange protective gear like masks, sanitizers, etc. The reverse transcription-polymerase chain reaction test i.e., RT-PCR test is the only test available to date to confirm the presence of the COVID-19 virus. Some researches show that CT scan images of the chest region are showing some signs of the early COVID-19 infection [41]. It is a quite inspiring finding because CT scan images can be used to isolate some suspected patients at an early phase and therefore, the drastic spread of this virus can be stopped to some extent. The CT scan images cannot replace the RT-PCR test because some false negatives are reported in [42,43]. The screening of the COVID-19 positive patients using the CT scan images are recommended in [44]. The presence of some prominent features like ground-glass opacities, crazy paving, etc. (which are given in Table 2) helps us to trace the initial presence of this infection, and image segmentation is an essential task to automate the screening process. Typically, the absence of properly annotated data makes the automated biomedical image analysis job difficult. These are the basic motivation behind proposing a novel radiological image segmentation approach SUFACSO (Superpixel based Fuzzy Artificial Cell Swarm Optimization). As the name suggests, the proposed approach is based on the superpixels and type 2 fuzzy systems where the type 2 fuzzy objective function is modified to incorporate the
advantages of superpixels to efficiently process a large amount of spatial information. The fuzzy objective function is optimized with the recently developed metaheuristic procedure i.e., artificial cell swarm optimization. The proposed method allows automated and efficient analysis of the CT scan images which is beneficial to enhance the computer-aided diagnostic systems to act as a tool against the COVID-19 virus.

To summarize, the major contributions are as follows: (1) A novel superpixel-based image segmentation technique is proposed that reduces the incurred computational cost for processing a high amount of spatial information, (2) Type-II fuzzy system is incorporated with the superpixel-based approach, (3) A recently developed metaheuristic procedure ACSO is further enhanced, (4) The conventional fitness function of the FCM clustering approach is enhanced to exploit the advantages of superpixel (5) The cluster centers are updated with the help of the proposed fuzzy ACSO approach.

The remaining article is prepared in the following way: Sections 2 and 3 describes the artificial cell swarm optimization method and the type 2 fuzzy clustering framework respectively. Sections 4 and 5 describe the proposed SUFACSO approach and the obtained results respectively. Section 6 discusses some of the relevant points and a brief conclusion is presented in Section 7.

2. A brief overview of the artificial cell swarm optimization procedure

This is a recently developed metaheuristic procedure that is inspired by the artificial cell division procedure. The artificial cell swarm optimization procedure mimics the artificial cells as the search agents. The actual artificial cell division approach [46] is slightly modified to design the optimization procedure. The incorporated modifications are listed below [47]:

Fig. 2. Dependency of the number of superpixels on the size of the disk structuring element (a)-(h) superpixel images obtained using the disk structuring element of size 3 to 10 respectively, (i) Size of the structuring element vs. the number of superpixels.
The artificial cells are not depending on the current state to participate in the cell division process.

ii The artificial cells only take part in the cell division process

The hierarchical tree structure is formed throughout the generations due to the artificial cell division process. Swarms of artificial cells are considered in the optimization process to take part in the artificial cell division process. No communication is allowed between any pair of artificial cells. Lifespan of the kth artificial cell at a certain timestamp \( t_s \) is an important parameter and it is directly dependent on the fitness value \( \text{fitness}_k \) as given in Eq. (1).

\[
LS_k^{t_s} \propto \text{fitness}_k \tag{1}
\]

A huge number of swarms can significantly increase the fitness evaluations and a small number of swarms can increase time to converge and therefore, is essential to decide the swarm count moderately. In this work, the swarm count is considered is a fixed parameter. One artificial cell can produce some new cells and the production of new cells occur at a certain distance which is inversely dependent on the fitness of the producer cell as expressed in Eq. (2).

\[
\begin{align*}
\text{dist}_k^{ts} & \propto \frac{1}{\text{fitness}_k}, \\
\text{dist}_k^{ts} & \propto LS_k^{t_s-1}
\end{align*}
\tag{2}
\]

The distance between the kth cell and any of the lth cell, which are produced from the same parent cell, must be same. Therefore, if a cell is near to global optima, then it can generate some other cells at a smaller distance and vice-versa. Smaller steps help to search the nearest portions of the global optima cautiously so that the global optima may not be missed accidentally. A cell does not have any effect on the population once its lifespan is over. This property helps to maintain the size of the population and prevents getting overpopulated. The successor cells of a cell can produce some other cells by the cell division process to maintain
the population. The life span of a cell can belong to the near-optimal area. The quality of a population is evaluated using the lambda function which is given in Eq. (3).

$$\lambda \left( \sum_k \text{fitness}^k \right), \ R \rightarrow [0, 1]$$

(3)

The tentative population at timestamp \( t_s + 1 \) can be determined using Eq. (4) where \( \psi \) denotes the productivity.

$$P (t_s + 1) = \psi \cdot \lambda \left( \sum_k \text{fitness}^k \right) \cdot P (t_s)$$

(4)

Algorithm 1 illustrates the artificial cell swarm optimization approach in brief [47].

3. Fuzzy C-means clustering based on type 2 fuzzy system

The proposed approach adopts the type 2 fuzzy logic-based clustering approach to effectively model and handle the random uncertainties. In most real-life applications, the uncertainty cannot be predicted in advance. A wide range of input types can produce random uncertainties. Hence, it is essential to cope up with the random uncertainties in real-life scenarios. The fuzzy C-means clustering approach is one of the widely used clustering
approaches which is suitable to various problems of different domains [48–51]. The main reason behind the increasing popularity of fuzzy systems is the suitability of this approach in different scenarios where the crisp clustering approaches do not perform well. A single point can be a member of more than one cluster at the same time with some membership values. The total sum
of all membership values for a certain point must be one. So, the value of the membership can be anything between 0 and 1. The dissimilarity function which is optimized by the fuzzy C-means clustering approach is given in Eq. (5).

\[
Obj_\psi = \sum_{k=1}^{n}\sum_{l=1}^{nClstrs} \mu_{kl}^\psi \|x_k - c_l\|^2, \text{ where } 1 \leq \psi < \infty
\]  

(5)

The value of the membership (\(\mu_{kl}^\psi\)) can be computed using Eq. (6) and \(\psi\) denotes the fuzzifier. The cluster centers can be updated using Eq. (7).

\[
\mu_{kl} = \frac{1}{\sum_{l=1}^{nClstrs} \left( \frac{|x_k - c_l|}{|x_k - c_l|} \right)^\frac{2}{\psi - 1}}
\]  

(6)

\[
c_l = \sum_{k=1}^{nPts} \mu_{kl}^\psi x_k \frac{\sum_{k=1}^{nPts} \mu_{kl}^\psi}{\sum_{k=1}^{nClstrs} \mu_{kl}^\psi}
\]  

(7)

The type 2 fuzzy logic systems use separate sets of membership values that are also fuzzy in nature. This approach allows efficient modeling of dynamic input uncertainties by providing additional degrees of freedom. In this work, the type 2 fuzzy logic-based clustering approach is adopted to overcome some of the common problems of type 1 fuzzy systems like noise sensitivity, relative membership values, etc., and also to handle uncertainties well [52]. It is essential to improve the outcome of the segmentation process. The uncertainty of a point must be decided depending on the membership value i.e., if a point has a membership value of 1 then its uncertainty will be certainly nil. So, a lower membership value indicates higher uncertainty and
Some of the basic reasons behind the adoption of type 2 fuzzy system in this work are listed below [53]:

a. A point with higher uncertainty has a lesser impact on the overall clustering process and vice-versa. It helps to achieve more realistic results.

b. Better noise handling capability can be achieved.

The membership value in type 2 fuzzy systems can be calculated using Eq. (8) and the cluster centers can be updated using Eq. (9). The proposed approach does not require Eq. (9) and can update the cluster centers. The artificial cell swarm optimization process will guide the proposed approach to determine the optimal cluster centers. The accuracy clustering process can be determined by a small threshold value $\varphi$. The type 2 fuzzy clustering system can be easily understood from algorithm 2 and the schematic diagram of the type 2 fuzzy system can be visualized from Fig. 1.

$$\dot{\mu}_{kl} = \mu_{kl} - \frac{1 - \mu_{kl}}{2}$$ (8)

$$\dot{c}_l = \frac{\sum_{k=1}^{nPts} \mu_{kl}^\psi x_k}{\sum_{k=1}^{nPts} \mu_{kl}^\psi}$$ (9)
Fig. 7. (continued).

Fig. 8. Performance comparison of different algorithms for different cluster validity indices (a) Davies–Bouldin, (b) Xie–Beni, (c) Dunn, and (d) $\beta$ index. In X-axis the number of clusters and in the Y-axis, the values of the corresponding validity index are plotted.
4. Proposed SUFACSO approach for CT scan image explication

4.1. Proposed method of superpixel computation

The ever-growing technology allows us to increase the quality of the image acquisition hardware. High-quality biomedical images can be acquired from various biomedical image acquisition devices and it is helpful in a precise analysis of the biomedical images. Automated biomedical image analysis devices are facing some challenges due to the increasing quality of biomedical images. A high amount of spatial information creates severe problems for automated and computer-aided diagnostic systems because medical diagnostic systems demand quick and accurate results. Image segmentation plays a vital role in many automated computer-aided image analysis systems. It is essential to generate precise reports within the stipulated amount of time to provide accurate treatment to the patients. To handle this situation effectively and to accelerate the screening process of the COVID-19 infection, a superpixel-based novel approach is proposed in this work to segment the CT scan images. Superpixels are useful to represent a set of pixels in a computation-friendly manner.

Different approaches can be found in the literature to find the superpixel image from an input image [54–56]. Some superpixel computation methods like mean shift [54] and watershed [56] produce irregular superpixels and some methods like SLIC [55] generate regular superpixels. Meanshift and watershed approaches are more useful due to the capability to generate irregular superpixels. The watershed approach is simpler to implement compared to the mean-shift approach but it is sensitive
| Reference/ Source | Method | Type of the biomedical image | Comments |
|-------------------|--------|------------------------------|----------|
| Jentzen et. al. [18] | Iterative thresholding | Positron Emission Tomography volumes | This approach is used to segment the PET volumes based on varying source-to-background (S/B) ratios which are collected from the phantom of a body. The calibrated source-to-background curves are used to determine the volume using the iterative thresholding procedure. One major drawback of the system is that it cannot effectively measure small volumes. |
| Wiemker et. al. [19] | Thresholding | CT scan images | This approach segments the CT scan images to easily interpret and study the lung nodules. This work proposes a divergence theorem and histogram-based CT image segmentation approach. This approach is can effectively and optimally isolate the lung nodules from the CT scan images. In this context, the optimality is defined in terms of the mean gradient of the iso-surface and the sphericity. |
| Asari et. al. [20] | Thresholding and differential region growing | Endoscopic images | This work is targeted to extract gastrointestinal lumen from the endoscopic images. This algorithm is consisting of two stages where the first stage employs a global thresholding approach and in the second phase, the differential region growing is used to extract the gastrointestinal lumen from the endoscopic images. The dynamic hill-clustering approach is used to ascertain the effectiveness of the segmentation and to look after the growth process. |
| Yu-qian et. al. [21] | Edge detection | CT scan images | This work is addressing the problem of edge detection in the presence of noise. Traditional gradient-based edge detection approaches are susceptible to noise and therefore, this approach proposes a novel approach to detect edges of the lung CT scan images using mathematical morphology. This approach is tested on the CT images which are corrupted with the salt-and-pepper noise and its efficiency is proved by comparing this approach with some of the other standard approaches. It is observed that this approach can efficiently reduce the effect of noise and also can generate precise edges. |
| Falcao et. al. [22] | Shortest-path based method | MRI images | This work is based on the computation of the shortest path using Dijkstra's algorithm. This approach is highly dependent on the user intervention to efficiently determine the segmented regions and to define the objects. This approach is found to be 3 to 15 times faster compared to manual tracing. This approach can be applied almost independently to the applications. One main problem associated with this method is the difficulties associated with the choice of slabs and orthogonal slices which has a serious impact on the efficiency of this approach. |
| Pan et. al. [23] | Edge detection | Cellular image | This work proposes a novel edge detection approach which is based on the bacterial foraging algorithm. The proposed approach addresses the problem of discontinuous edges and dependency on the initialization which are associated with the traditional edge detection approaches. In this work, the intensity of the gradient images is modeled as the concentration of the nutrients and the property of the bacteria Escherichia coli. The edges are highlighted as the paths of the bacteria. Although this approach performs well and comparative study shows the effectiveness of the proposed approach still, one problem of this approach is not very robust to noise. Noise can lead to crumpled edges. This approach is not also suitable to handle overlapped cells. |
| Ji et. al. [24] | Fuzzy C-means clustering | Synthetic, MR images, natural images | This work proposes a modification to the traditional fuzzy C-means clustering by addressing some of the problems. Traditional fuzzy C-means clustering approach does not consider the spatial information and less robust to noise. This work proposes a modification which is known as the weighted image patch-based FCM. In this work, pixels are replaced with the weighted patches which is helpful to incorporate spatial information in the segmentation process. It is helpful to increase the reliability of the overall segmentation process but it also increases the computational overhead drastically. |

(continued on next page)
A controlling parameter of structuring element is decided as per the range of the size applying more than one structuring elements where the number of the image. Therefore, the pointwise maximum value is computed for each structuring element which are of various sizes, depending on the image under consideration. Practically, it is not possible to determine different structuring elements which are of various sizes, depending on the image. Therefore, the pointwise maximum value is computed as per the range of the size controlling parameter \( \phi \) i.e., \( [\phi_1, \phi_2] \in \mathbb{N}^+ \) and \( \phi_1 \leq \phi \leq \phi_2 \).

The number of superpixels is inversely dependent on the size of the structuring elements. It can be easily understood in Figs. 2 and 3 and, Figs. 2(i) and 3(i) graphically depicts this fact. The image considered in these two figures is the Test01 image [58] (please refer to Table 3). Figs. 2(i) and 3(i) plots the count of the superpixel in the \( y \)-axis and the size of the structuring element is plotted in the \( x \)-axis.

\[
\hat{R}_{\text{img}}^\phi (\text{img}', \phi, \phi_1) = \max \left\{ R_{\text{img}}^\phi (\text{img}')_{\phi_1}, R_{\text{img}}^\phi (\text{img}')_{\phi_1+1}, \ldots, R_{\text{img}}^\phi (\text{img}')_{\phi_2} \right\}
\]  

A very small lower bound is not desirable because it will produce very small regions and some essential edge information can be lost. A small threshold value \( \psi \) is used to control the error rate and the upper threshold value as given in Eq. (17). A higher value of \( \psi \) indicates a higher error rate but, a smaller upper bound that helps to achieve lesser computational overhead. So, the threshold value can be adjusted as per the requirement and depending on the available resources.

\[
\begin{align*}
\hat{R}_{\text{img}}^\phi (\text{img}', \phi, \phi_1) - \hat{R}_{\text{img}}^\phi (\text{img}', \phi, \phi_1 + 1) \leq \psi
\end{align*}
\]  

### 4.2. Proposed superpixel coupled fuzzy ACSO approach-based segmentation

The conventional fuzzy C-means clustering approach often overlooks some important spatial information that can be costly in terms of the segmentation performance. Some approaches try to solve this problem by considering and blending some local spatial information in the objective function but it increases the computational cost and therefore not suitable on many occasions. Superpixels can help in this context by over-segmenting an image in many small, perceptually uniform, and homogeneous regions. In this work, the CT images are first processed to determine the segmented image by finding the optimal clusters. As discussed earlier, the type 2 fuzzy system is used to perform the segmentation. The fuzzy objective function which is given in

| S. Chakraborty and K. Mali | Applied Soft Computing 119 (2022) 108528 |
|---------------------------|-----------------------------------------|
| **Table 1 (continued)**   |                                         |
| Reference/ Source         | Method                                   | Type of the biomedical image | Comments                                                                 |
| Agrawal et. al. [25]      | Optimum boundary point detection         | MR images                    | This approach is devoted to segment the intracranial region from the magnetic resonance images. This work proposes a novel hybrid approach which is based on the genetic algorithm and the bacterial foraging algorithm. The combination of these two approaches is used to optimize the objective function of the fuzzy c-means clustering. The final cluster centers are obtained using a method called optimum boundary point detection. This approach cannot determine the optimal number of clusters automatically and produces inaccurate results if the predefined clusters and the actual number of clusters differ. |
| Chaia [26]                | Fuzzy C-means clustering                  | CT images                    | This work proposes a new approach to segment CT images. This approach is based on intuitionistic fuzzy set theory and it is known as the intuitionistic fuzzy C means clustering. In this work, a novel objective function which is known as intuitionistic fuzzy entropy is incorporated with the traditional fuzzy C-means clustering. This approach is applied to different CT scan images to prove its efficiency. |
| Miao et. al. [27]         | Dictionary learning and Improved fuzzy C-means clustering | Synthetic, MRI, CT Scan | In this work, a noise-resistant version of the fuzzy c-means clustering algorithm is proposed and applied to segment the images. This approach can be divided into two phases where the first phase incorporates a dictionary learning method to handle the noise. In the second phase, this dictionary learning approach is hybridized with the Improved fuzzy c-means clustering approach. The proposed approach is not efficient for medical images with inhomogeneous intensity distribution. |

Table 2

Significant properties which are found in the CT scan images of the chest region of COVID-19 positive patients [45].

| Finding                                      | Percentage of the observed samples |
|----------------------------------------------|-----------------------------------|
| ground-glass opacities (GGO)                 | 100%                              |
| Multilobe and posterior involvement          | 93%                               |
| Bilateral pneumonia                          | 91%                               |
| Subsegmental vessel enlargement (3 mm)       | 85%                               |

are given in Eqs. (10) and (11) respectively.

\[
R_{\text{img}}^\phi (\text{img}') = R^\phi (R^\phi) \tag{10}
\]

\[
R_{\text{img}}^\phi (\text{img}') = R^\phi (R^\phi) \tag{11}
\]

In these equations, \( \gamma \) and \( \lambda \) denotes the morphological dilation and the erosion respectively which are expressed in Eqs. (12) and (13).

\[
\lambda_{\text{img}} (\text{img}') = \lambda (\lambda^{\gamma-1} (\text{img}')) \lor \text{img}' \tag{12}
\]

\[
\gamma_{\text{img}} (\text{img}') = \gamma (\gamma^{\lambda-1} (\text{img}')) \land \text{img}' \tag{13}
\]

In the above equations, \( \text{img} \) and \( \text{img}' \) denotes the original image and the marker image and the \( \text{img}' \) can be expressed by Eqs. (14) and (15). \( \lor \) and \( \land \) are the two operators to compute the pointwise maximum and the minimum values.

\[
\text{img}' = \gamma_{\text{ref}} (\text{img}) \tag{14}
\]

\[
\text{img}'' = \lambda_{\text{ref}} (\text{img}) \tag{15}
\]

Here, \( \text{se} \) is the structuring element and it is an important parameter that controls the segmented outcome. The size of the structuring element is subjective and depends on the image under consideration. Practically, it is not possible to determine different structuring elements which are of various sizes, depending on the image. Therefore, the pointwise maximum value is computed (using Eq. (16)) from the gradient images, which are generated by applying more than one structuring elements where the number of structuring elements is decided as per the range of the size controlling parameter \( \phi \) i.e., \( [\phi_1, \phi_2] \in \mathbb{N}^+ \) and \( \phi_1 \leq \phi \leq \phi_2 \).
Algorithm 3: The proposed SUFACSO approach

**Input:** The input image.

**Output:** Optimal segmented output image

1. Apply the method described in [57] to generate the gradient image.
2. Determine the superpixel from the gradient image using equation 11 and 12.
3. Randomly initialize the cluster centers $C_{n} = i_{t} + \text{rand}(0, 1) \cdot (i_{h} - i_{l})$ where,
   $k = \{1, 2, 3, \ldots, nClstrs\}$ and $i_{h}$ and $i_{l}$ denotes the highest and the lowest intensity values.
4. Apply equation 20 to compute the value of the modified objective function for each solution (i.e., determine the value of the fuzzy objective function for each artificial cell).
5. For each artificial cell, allot the value of the $LS$.
6. $itrCount \leftarrow 1$
7. Repeat while $itrCount \leq \text{max} \text{itrCount or !stoppingCriteria}$
   8. Compute the fitness of each cell.
   9. Generate the successor cells using the ACD process by following the successor distance
   10. Assign the value of $LS$ for each cell.
   11. Modify the $LS$ value of the parent and also update the successor distance.
   12. Apply equation 22 to determine the fuzzy membership for each point.
   end while
13. Construct the segmented image as per the computed optimal cluster centers
14. Return the segmented output image

Eq. (5), is optimized by the artificial cell swarm optimization algorithm. To incorporate the advantages of the superpixel, it is necessary to modify the fuzzy objective function. The fuzzy objective function which is given in Eq. (5) deals with the pixel-wise spatial information. The concept of superpixel represents a group of a pixel using a single value $Z_{w}$, as given in Eq. (18).

$$Z_{w} = \frac{1}{\text{cntPix}_{w}} \sum_{x \in B_{m}} \text{pix}_{x}$$ (18)

$$\text{Obj}_{\psi} = \sum_{t=1}^{\text{nClstrs}} \sum_{u=1}^{\text{cntPix}_{w}} \mu_{tu} \| Z_{t} - c_{u} \|^{2}, \text{ where } 1 \leq \psi < \infty$$ (19)

The value of the membership $\mu_{tu}$ can be computed using Eq. (20) and the corresponding type 2 fuzzy membership value can be computed using Eq. (21).

$$\mu_{tu} = \frac{1}{\sum_{v=1}^{\text{nClstrs}} \left( \frac{\| Z_{t} - c_{v} \|}{\| Z_{t} - c_{u} \|} \right)^{2/\psi}}$$ (20)

$$\dot{\mu}_{tu} = \mu_{tu} - \frac{1 - \mu_{tu}}{2}$$ (21)

The cluster centers can be updated and guided by the artificial cell swarm optimization and therefore, no equation is required to compute the updated positions of the cluster center. This approach is not dependent on the selection of the initial cluster centers. Cluster centers are initialized in a random order $C_{n} = i_{t} + \text{rand}(0, 1) \cdot (i_{h} - i_{l})$ where, $k = \{1, 2, 3, \ldots, nClstrs\}$ and $i_{h}$ and $i_{l}$ denotes the highest and the lowest intensity values. The proposed procedure is given in algorithm 3 and the schematic flow diagram is given in Fig. 4.

5. Experimental results

The performance evaluation and comparison of the proposed SUFACSO approach are presented in this section. As discussed earlier, the properly annotated ground truth segmented images may not be available always, and therefore, some standard intrinsic cluster evaluation methods are used here to evaluate the proposed approach quantitatively. Davies–Bouldin index [59], Xie–Beni index [60], Dunn index [61] and $\beta$ index [62] are some of the popular and frequently used intrinsic cluster validity indices which are used in this work for the evaluation purpose and these indices are defined in Eqs. (22) to (25) respectively.

$$DBIndex = \frac{1}{C} \sum_{u=1}^{C} \max_{i \neq k} \left( \frac{d_{u} (a_{i}) + d_{u} (a_{k})}{d_{u} (a_{i}, a_{k})} \right), \text{ i \neq k}$$ (22)

$$XBIndex = \frac{1}{C} \sum_{u=1}^{C} \sum_{q=1}^{C} \mu_{pq}^{2} \left( \frac{d_{min} \| C_{p} - X_{q} \|^{2}}{d_{max} \| C_{r} - C_{q} \|^{2}} \right)$$ (23)

$$DI_{n} = \min_{1 \leq i \leq n} \left( \min_{1 \leq j \leq n, j \neq i} \left( \frac{\text{dist} (c_{i}, c_{j})}{\max_{1 \leq k \leq n} \| C_{k} \|} \right) \right), \text{ } \gamma \text{ is the mean dist. of the pair wise clusters}$$ (24)

$$\beta = \frac{\sum_{u=1}^{C} \sum_{v=1}^{C} \mu_{uv} \left( I_{uv} - \chi \right)^{2}}{\sum_{u=1}^{C} \sum_{v=1}^{C} \left( I_{uv} - \frac{1}{\mu_{u}} \sum_{p=1}^{C} \mu_{vp} \left( I_{pu} \right)^{2} \right)}$$ (25)

5.1. Dataset description

200 CT scan images of the chest region are collected from the COVID-19 positive patients from different geographic regions. The proposed methods are applied to the 200 images and the test results are demonstrated with the 10 CT scan images that are randomly selected which are obtained from different countries of the world. Table 3 gives a brief overview of the test images and the test images along with their histograms are given in Fig. 5.
5.2. Experimental results

The experiments are performed in the MatLab R2014a on a computer that is equipped with an Intel i3 processor and 4 GB main memory. The proposed method is compared with some metaheuristic optimization-based image segmentation approaches like modified genetic algorithm [67], modified PSO [68], improved bat algorithm [69] and modified cuckoo search method [70] in both qualitative and quantitative manner. The visual comparison is presented in Fig. 6 where the Test01 image is considered. The segmented output images which are obtained by applying the proposed SUFACSO approach, are reported in Fig. 7. The quantitative comparative study is reported in Tables 4 to 7 for the Davies–Bouldin index [59], Xie–Beni index [60], Dunn index [61] and $\beta$ index respectively. The acceptable values are highlighted in boldface. The comparisons and evaluations are performed for different numbers of clusters.

From the qualitative and quantitative results, it can be observed that the proposed SUFACSO approach outperforms some state-of-the-art works and can produce realistic outputs that are certainly helpful for the interpretation of the real-life CT scan images and therefore, this approach can be helpful for the early screening purposes. At the end of each table, the average performance of the five approaches is reported which is beneficial to understand the overall performance of these methods for the different number of clusters and different cluster validity indices. In the case of average, the column-wise optimal values are highlighted instead of highlighting the row-wise optimal values. The row-wise highlighted values talk about the performance of the individual algorithm for the different number of clusters whereas the column-wise highlighted values help to understand the performance of the individual algorithms. It can be observed that the proposed approach outperforms other approaches for most of the number of clusters as well as for most of the validity indices. For example, on a total of 16 occasions (i.e., 4 validity indices x 4 different number of clusters), the proposed approach is found to perform better 11 times. These comparative results are graphically presented in Fig. 8.

The experiments are carried out for the different numbers of clusters. A particular approach may perform well for a particular cluster count. That is why the average values of all experiments are reported at the end of each table for better interpretation. It can be observed that the proposed approach can optimize different objective functions effectively.

Actually, the experiments are carried out on 200 CT images (in the first phase) and 100 CT images (in the second phase). It is already mentioned in Section 5.1. Results that are obtained from all images are not possible to report in this stipulated amount of space. Therefore, only some results that are obtained from some selected images are reported.

Apart from these tests, the proposed approach is also compared with some of the active contour models based on some standard parameters like accuracy, precision, and recall. This comparison is performed by using the database that is available at [71]. This dataset contains 100 CT scan images with dimensions $512 \times 512$. This dataset is created by collecting sample images from 49 patients with age range 32–86 years. The obtained average results are reported in Table 8.

| Test | Source | Gender | Age | Observed properties | Comments |
|------|--------|--------|-----|---------------------|----------|
| Test01 | Coronal Axial | [63] | F | 50 | ground-glass opacities (GGO) | Case courtesy of Dr Bahman Rasuli, Radiopaedia.org, rID: 75329 |
| Test03 | Axial Coronal | [64] | M | 75 | ground-glass opacities (GGO) crazy paving enlarged mediastinal lymph nodes | Case courtesy of Dr Fabio Macori, Radiopaedia.org, rID: 74867 |
| Test05 | Axial Coronal Sagittal | [65] | F | 70 | ground-glass opacities (GGO) crazy paving air space consolidation | Case courtesy of Dr Ammar Haouimi, Radiopaedia.org, rID: 75665 |
| Test08 | Sagittal Axial Coronal | [66] | M | 50 | ground-glass opacities (GGO) | Case courtesy of Dr Ammar Haouimi, Radiopaedia.org, rID: 76295 |

5.3. Study of the convergence rate

The rate of convergence is an important parameter to be studied. The performance evaluation remains incomplete without studying and comparing the convergence of different algorithms. The convergence analysis gives a clear view of the comparative performance of different algorithms for the different numbers of clusters. The graphical analysis of the convergence is presented in this subsection using the image Test01 for the Dunn index. In Fig. 9, five different plots are given for five different methods. In a single plot, four separate curves are indicating four different clusters. These curves show that the proposed approach can efficiently segment the images for a higher number of clusters. Moreover, the proposed approach also outperforms some other methods in terms of convergence besides quantitative and qualitative performance.

5.4. Analysis of the complexity

The time complexity is an important aspect that is to be analyzed. From the detailed discussion of the proposed approach, it can be noticed that the proposed approach can be viewed as a two-phase procedure where the watershed-based computation approach is used to determine the superpixel image from the underlying image in the first phase and the optimal segmented outcome is computed in the second phase. The task of optimization is performed using the proposed fuzzy ACSO approach. The gradient information of an image is used to avoid the noise sensitivity of the water-shed based superpixel computation process.

The watershed-based technique is a simple method to compute the superpixel and the implementation follows linear complexity [56]. It is quite inspiring and lucrative to adopt this approach on different occasions. In the optimization part, the fuzzy objective function is optimized by using the proposed fuzzy ACSO method. The ACSO approach is an effective and efficient approach that can be executed in linear time [47]. So, the proposed approach is efficient enough and can be effectively used in various real-life problem-solving scenarios.

The proposed SUFACSO approach is basically an unsupervised clustering approach that is used for image segmentation purposes. This approach can effectively process high-quality images with the help of the proposed superpixel-based approach that is an essential quality for the real-life application of an image segmentation approach. This approach removes the dependency...
of choice of the initial cluster centers as well as the ACSO approach determines the optimal cluster centers by optimizing some validity indices. These advantages motivate us to apply the proposed approach to automatically segment the radiological images that will be certainly helpful in diagnosing some symptoms of COVID-19. The experimental outcomes show the efficiency of the proposed approach. Under this pandemic environment, this work is designed hoping that it can help physicians and other domain experts to some extent in the early diagnosis of the disease. Early diagnosis can prevent the drastic spread of this highly infectious virus.

The quantitative outcomes of the proposed SUFACSO approach are useful to assess the comparative performance. Quantitative results do not have any direct implications in real-life diagnosis. The segmented outcomes are useful in the diagnosis process. Physicians can investigate the segmented outcomes to find some prominent and common features as mentioned in Table 2. The segmented images will be helpful in the easy interpretation of the radiological images.

The proposed SUFACSO approach is an efficient image segmentation approach that can effectively segment the radiological images that highly useful in the easy interpretation of these
images. The presence of the COVID-19 virus can only be confirmed with the help of some standard tests and one of the most popular and gold-standard tests is the RT-PCR test. Typically, the test reports of the RT-PCR tests are generated within 2–4 days. There is a high possibility that a suspected patient can spread the disease in the community completely unwillingly. The proposed approach can reduce this chance because an initial screening can be performed by the physicians comfortable with the help of the proposed SUFACSO approach. It is worth mentioning here that the proposed approach is neither a replacement of the RT-PCR test nor it can confirm the presence of the virus accurately. However, this approach can be helpful in an initial screening at an early stage that will restrict the spread of this highly infectious virus by separating suspected patients from the rest of the community.

6. Discussion

6.1. Threats to validity

The obtained results indicate that the proposed approach is suitable for real-life scenarios and also performs efficiently. This approach can be easily adapted for the automated screening purposes of the COVID-19 infected patients. It is assumed the quality of the CT scan images is considerably high and the performance
knowledge of the authors, there is no publicly available manually.

Results are quite promising and encouraging. From the best of the

performance of the proposed approach is not verified against the presence of

noise. It will be interesting to study the proposed approach in

the presence of noise. The scalability of the proposed approach
to different types of biomedical images can be explored in future

studies. Missing manual annotations can jeopardize the generalizability of the proposed work. On the other hand, the obtained results are quite promising and encouraging. From the best of the knowledge of the authors, there is no publicly available manually

annotated dataset for the chest CT scan images of the COVID-19

positive cases.

### 6.2. Limitations

Although the proposed approach is efficient enough to segment the CT scan images automatically and produces realistic segmented outcomes, some misclassifications can be observed in this proposed approach that can be addressed in the subsequent works. One important drawback of the proposed approach is that it cannot automatically determine the number of clusters and it can be overcome in future works. Automated estimation of the clusters can make this approach more realistic, robust, and application friendly. The proposed method can handle only a single objective at a time. Therefore, the proposed
The proposed approach cannot automatically differentiate between different types of diseases. For example, the approach cannot take any pre-trained model. The proposed approach can effectively segment the radiological images that are collected from different patients. It neither uses any training dataset nor uses any additional CT images of COVID-19 infection as well as on some standard dataset of the biomedical images.

The proposed SUFACSO approach is an unsupervised segmentation approach. It neither uses any training dataset nor uses any pre-trained model. The proposed approach can effectively segment the radiological images that are collected from different patients, i.e., not only COVID-19 infected samples but samples collected from patients with other infections as well as normal patients. It is to be clarified that this approach cannot take any decision about the type of disease automatically. For example, the proposed approach cannot automatically differentiate between
COVID-19 related lung images and other lung diseases. This approach aims to help physicians in early and quick interpretation of the radiological images and diagnosis of the diseases without any manual delineations.

7. Conclusion

This article proposes a novel, simple and elegant solution that uses some of the important features of the chest CT scan images to screen the COVID-19 suspected patients easily and at an early phase which can be considered as an effective tool to reduce the drastic spread of this virus. From Fig. 8, it can be observed that the proposed approach works well in most situations and outperforms most of the other standard approaches. Both qualitative and quantitative study produces some satisfactory results which help to make the proposed approach trustworthy so that it can be reliably adapted in the real-world scenarios. From Fig. 9, it can be observed that the proposed approach performs well in terms of convergence. The proposed approach initially performs a superpixel-based clustering using the proposed superpixel computation method which significantly reduces the computational overhead for the further clustering process by reducing a large amount of spatial information. Therefore, radiological images can be conveniently explicated with the application of the proposed method and the proposed approach is also helpful in the easy interpretation of the radiological images. The proposed work neither claims that the suggested approach is cent percent accurate in determining the COVID-19 infection nor claims that it can be a replacement of the RT-PCR test but, the proposed method can help detect some common characteristics from the CT scan images, that can help to isolate some suspected patients from the rest of the community. The proposed approach is helpful for the early screening of the COVID-19 besides being a significant contribution to the image segmentation literature.

CRediT authorship contribution statement

Shouvik Chakraborty: Conceptualization, Methodology, Software development, Investigation. Kalyani Mali: Formal analysis, Resources, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to express their gratitude and thank the editors, anonymous reviewers, and referees for their valuable comments and suggestions which are helpful in further improvement of this research work.

Appendix. Experimental setup and environment

Hardware setup

The experiments are performed in a laptop PC which is equipped with 4 GB of RAM, 1 gigabyte of dedicated graphics memory, Intel Core i3-3217U processor with 1.8 GHz clock speed. The dedicated graphics memory is not utilized for any kind of processing purposes.

Software setup

The system in which the experiments are carried out is equipped with the Microsoft Windows 7 (64 bit) operating system. The proposed SUFACSO approach is coded in the Matlab R2014a environment. It is not at all essential to use the Matlab environment to implement the proposed approach. We have chosen Matlab due to the availability of some inbuilt functions which are helpful to reduce the coding complexity. Still, any other languages or platforms can be used to implement the same. The graphs are also generated in the Matlab environment where the ‘best’ position is selected to accommodate the legend without interfering the actual plot.

Information about the image data and segmented output

The proposed SUFACSO approach is applied to the CT scan images which are collected from the chest region of the COVID-19 infected patients. It is assumed that there are no manual annotations available and the proposed approach is capable to process the images without having any prior knowledge. The proposed SUFACSO approach produces the optimal segmented image by computing the optimal cluster centers. The final segmented images are constructed by assigning the superpixel to their corresponding cluster centers. These segmented images are helpful to interpret different features from these radiological images.

References

[1] T.H. Kim, K.M. Lee, S.U. Lee, Learning full pairwise affinities for spectral segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 1690–1703, http://dx.doi.org/10.1109/TPAMI.2012.237.
[2] S. Chakraborty, K. Mali, An overview of biomedical image analysis from the deep learning perspective, in: S. Chakraborty, K. Mali (Eds.), in: Appl. Adv. Mach. Intell. Comput. Vis. Object Recognit. Emerg. Res. Oppor., IGI Global, 2020, http://dx.doi.org/10.4018/978-1-7998-2736-8.ch008.
[3] S. Chakraborty, M. Roy, S. Hore, A study on different edge detection techniques in digital image processing, in: Feature. Detect. Motion Detect. Video Process, IGI Global, 2016, pp. 100–122, http://dx.doi.org/10.4018/978-1-5225-1025-3.ch005.
[4] M.W. Libbrecht, W.S. Noble, Machine learning applications in genetics and genomics, Nat. Rev. Genet. 16 (2015) 321–332, http://dx.doi.org/10.1038/nrg3920.
[5] C. Tang, L. Zhang, A. Zhang, M. Ramanathan, Interelated two-way clustering: An unsupervised approach for gene expression data analysis, in: Proc. – 2nd Ann. IEEE Int. Symp. Bioinforma. Bioeng. BIBE 2001, 2001, pp. 41–48, http://dx.doi.org/10.1109/BIBE.2001.974410.
[6] S. Chakraborty, An advanced approach to detect edges of digital images for image segmentation, in: S. Chakraborty, K. Mali (Eds.), in: Appl. Adv. Mach. Intell. Comput. Vis. Object Recognit. Emerg. Res. Oppor., IGI Global, 2020, http://dx.doi.org/10.4018/978-1-7998-2736-8.ch004.
[7] S. Chakraborty, K. Mali, SUFMACS: A machine learning-based robust image segmentation framework for covid-19 radiological image interpretation, Expert Syst. Appl. (2021) 115069, http://dx.doi.org/10.1016/j.eswa.2021.115069.
[8] S. Chakraborty, K. Mali, Application of multiobjective optimization techniques in biomedical image segmentation—A study, in: Multi-Objective Optim, Springer Singapore, Singapore, 2018, pp. 181–194, http://dx.doi.org/10.1007/978-981-13-1471-1_8.
[9] S. Chakraborty, S. Chatterjee, K. Mali, An optimized intelligent dermatologic disease classification framework based on Iot, in: Adv. Intell. Syst. Comput. Springer, 2020, pp. 131–151, http://dx.doi.org/10.1007/978-981-15-4288-6_9.
[10] S. Hore, S. Chakraborty, S. Chatterjee, N. Dey, A.S. Ashour, L. Van Chung, D.-N. Le, An integrated interactive technique for image segmentation using stack based seeded region growing and thresholding, Int. J. Electr. Comput. Eng. 6 (2016) http://dx.doi.org/10.11591/ijece.v6i6.11801.
[11] S. Chakraborty, A. Raman, S. Sen, K. Mali, S. Chatterjee, H. Hachimi, Contrast optimization using elitist metaheuristic optimization and gradient approximation for biomedical image enhancement, in: 2019 Amyt Int. Conf. Artif. Intell. Intell., IEEE, 2019, pp. 712–717, http://dx.doi.org/10.1109/AICAI.2019.8701367.
[12] S. Chakraborty, K. Mali, S. Chatterjee, S. Banerjee, K. Roy, K. Deb, S. Sarkar, N. Prasad, An integrated method for automated biomedical image segmentation, in: 2017 4th Int. Conf. Opto-Electronics Appl. Opt. Optroinix 2017, 2018, http://dx.doi.org/10.1109/OPTRONIX.2017.8349978.
[13] P. Mesejo, Ó. Ibáñez, Ó. Cordón, S. Cagnoni, A survey on image segmentation using metaheuristic-based deformable models: State of the art and critical analysis, Appl. Soft Comput. J. 44 (2016) 1–29, http://dx.doi.org/10.1016/j.asoc.2016.03.004.
[14] X. Chen, L. Pan, A survey of graph cuts/graph search based medical image segmentation, IEEE Rev. Biomed. Eng. 11 (2018) 112–124, http://dx.doi.org/10.1109/RBME.2018.2798701.
[63] COVID-19 Pneumonia | radiology case | radiopaedia.org, 2020, (n.d.). https://radiopaedia.org/cases/covid-19-pneumonia-29 (accessed May 6, 2020).
[64] COVID-19 Pneumonia | radiology case | radiopaedia.org, 2020, (n.d.). https://radiopaedia.org/cases/covid-19-pneumonia-12 (accessed June 10, 2020).
[65] COVID-19 Pneumonia | radiology case | radiopaedia.org, 2020, (n.d.). https://radiopaedia.org/cases/covid-19-pneumonia-63 (accessed May 31, 2020).
[66] COVID-19 Pneumonia | radiology case | radiopaedia.org, 2020, (n.d.). https://radiopaedia.org/cases/covid-19-pneumonia-88 (accessed June 10, 2020).
[67] H.Z. Jia, A.Y.C. Nee, J.Y.H. Fuh, Y.F. Zhang, A modified genetic algorithm for distributed scheduling problems, J. Intell. Manuf. 14 (2003) 351–362, http://dx.doi.org/10.1023/A:1024653810491.
[68] B.F. Moghaddam, R. Ruiz, S.J. Sadjadi, Vehicle routing problem with uncertain demands: An advanced particle swarm algorithm, Comput. Ind. Eng. 62 (2012) 306–317, http://dx.doi.org/10.1016/j.cie.2011.10.001.
[69] X. Cai, X.Z. Gao, Y. Xue, Improved bat algorithm with optimal forage strategy and random disturbance strategy, Int. J. Bio-Inspired Comput. 8 (2016) 205–214, http://dx.doi.org/10.1504/IJBIC.2016.078666.
[70] S. Chakraborty, S. Chatterjee, N. Dey, A.S. Ashour, A.S. Ashour, F. Shi, K. Mali, Modified cuckoo search algorithm in microscopic image segmentation of hippocampus, Microsc. Res. Tech. (2017) 1–22, http://dx.doi.org/10.1002/jemt.22900.
[71] COVID-19 - Medical segmentation, 2021, (n.d.). http://medicalsegmentation.com/covid19/ (accessed October 20, 2021).