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.
Ambiguous D-means fusion clustering algorithm based on ambiguous set theory: Special application in clustering of CT scan images of COVID-19

Pritpal Singh a,*,1, Surya Sekhar Bose b,1

a Institute of Theoretical Physics, Jagiellonian University, ul.Łojasiewicza 11, Kraków 30-348, Poland
b Department of Mathematics, Madras Institute of Technology, MIT Rd, Radha Nagar, Chromepet, Chennai, Tamil Nadu 600044, India

A R T I C L E    I N F O
Article history:
Received 3 April 2021
Received in revised form 17 August 2021
Accepted 20 August 2021
Available online 26 August 2021

Keywords:
Coronavirus disease 2019 (COVID-19)
Computed tomography (CT)
Ambiguous set theory
Ambiguous D-means fusion clustering algorithm (ADMFCA)
Clustering

A B S T R A C T
Coronavirus Disease 2019 (COVID-19) has been considered one of the most critical diseases of the 21st century. Only early detection can aid in the prevention of personal transmission of the disease. Recent scientific research reports indicate that computed tomography (CT) images of COVID-19 patients exhibit acute infections and lung abnormalities. However, analyzing these CT scan images is very difficult because of the presence of noise and low-resolution. Therefore, this study suggests the development of a new early detection method to detect abnormalities in chest CT scan images of COVID-19 patients. By this motivation, a novel image clustering algorithm, called ambiguous D-means fusion clustering algorithm (ADMFCA), is introduced in this study. This algorithm is based on the newly proposed ambiguous set theory and associated concepts. The ambiguous set is used in the proposed technique to characterize the ambiguity associated with grayscale values of pixels as true, false, true-ambiguous and false-ambiguous. The proposed algorithm performs the clustering operation on the CT scan images based on the entropies of different grayscale values. Finally, a final outcome image is obtained from the clustered images by image fusion operation. The experiment is carried out on 40 different CT scan images of COVID-19 patients. The clustered images obtained by the proposed algorithm are compared to five well-known clustering methods. The comparative study based on statistical metrics shows that the proposed ADMFCA is more efficient than the five existing clustering methods.

1. Introduction

In December 2019, the first Coronavirus Disease 2019 (COVID-19) outbreak has been discovered in Wuhan, China [1]. This disease is caused by a novel virus, called Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) [2]. According to the World Health Organization (WHO) [3], this disease has spread around the world and reported 200,840,180 confirmed cases along with 4,265,903 deaths by the end of 6th August, 2021. Now, this disease has seriously affected the health, economic and social systems of advanced and emerging countries [4]. Therefore, the WHO has declared this disease as one of the deadliest pandemics of all time. However, the pandemic situation caused by this disease has yet to be taken seriously [5]. Due to this reason, several countries, including the United States, Brazil, India and Italy, have been seriously affected by this virus [6,7]. As a result, several research groups are collaborating in the development of strategies, vaccines and new approaches to address the pandemic [8].

COVID-19 is associated with serious respiratory symptoms that cause death in most cases [9]. This disease is observed as pneumonia, the infection of which is very contagious from one to the other [10]. Digital imaging techniques, such as X-ray [11] and computed tomography (CT) [12] have contributed significantly to the diagnosis of this disease. Computer scientists have been actively involved in developing methods for analyzing these images based on machine learning. To develop such methods, they mainly use the convolutional neural network (CNN), which is a form of deep neural network. Some of the CNN based methods devised by the researchers are summarized in Table 1.

The methodologies [11–19] discussed in Table 1 focus primarily on classifying X-ray or CT scan images in terms of infection and non-infection. Moreover, such approaches cannot identify the infection in the lungs. However, these studies indicate that X-ray and CT scan images of COVID-19 can be useful in determining the severity of lung infection. But, there are some inherent drawbacks.
Machine-learning based methods for X-ray and CT scan image analysis of COVID-19 patients.

| Article | Method developed | Image used | Objective |
|---------|------------------|------------|-----------|
| Ozturk et al. [13] | DarkNet model with 17 convolutional layers | X-ray | Binary (COVID-19 and No-finding) and multiclass (COVID-19, No-finding and Pneumonia) classification |
| Vaid et al. [14] | CNN with transfer learning | X-ray | Structural abnormalities and disease classification |
| Apostolopoulos and Mpesiana [11] | CNN with transfer learning | X-ray | Classification of abnormal X-ray images |
| Toraman et al. [15] | CapsNet model with capsule networks | X-ray | Binary (COVID-19 and No-finding) and multiclass (COVID-19, No-finding and Pneumonia) classification |
| Nour et al. [16] | CNN with k-nearest neighbor, support vector machine (SVM) and decision tree | X-ray | Classification of infection |
| Ardakani et al. [12] | CNN with machine learning | CT | Classification of positive COVID-19 cases |
| Kang et al. [17] | Multi-view representation learning | CT | Extraction of multiple features with different views |
| Wang et al. [16] | Weakly-supervised deep learning | CT | COVID-19 classification and lesion localization |
| Varela-Santos and Melin [19] | Feed-forward with CNN | X-ray | Classification of positive COVID-19 cases |

associated with X-ray and CT scan images of COVID-19, which are described below:

**Drawback 1**: The grayscale values in these images indicate some similarities between COVID-19 infection and other forms of pneumonia, which complicates decision-making.

**Drawback 2**: Large dimensions of these images increase the computational complexity during their processing.

**Drawback 3**: These images are available in multiple slices, but only a few of them indicate infection in the lungs.

**Drawback 4**: These images available with low-resolution and contain dense dark pixels, making it difficult to distinguish the infected lung regions.

The drawbacks mentioned in (1)–(4) can be overcome by various image processing techniques that help in detecting abnormalities in X-ray and CT scan images. Data clustering is one such method, whose main objective is to make a group of similar and dissimilar objects based on certain distance criteria [2,20,21]. However, clustering an image is one of the most difficult tasks in image processing, because in most cases it is difficult to form clusters of pixels by distinguishing similar grayscale values from dissimilar grayscale values. In the case of an image $K$, clustering is a means of partitioning it into non-overlapping regions of clusters $K_1, K_2, \ldots, K_n$, such that:

$$\bigcap_{i=1}^{n} K_i = \emptyset$$

$$\bigcup_{i=1}^{n} K_i = K$$

Based on the developments in clustering algorithms, they can be classified into two categories: (a) crisp clustering and (b) soft clustering [22]. Crisp and soft clusterings differ widely in their approach to assigning group members. In crisp clustering, the grayscale values of an image belong exclusively to a fixed group, i.e., their assignment is completely binary. That is, they completely belong to a cluster (true) or not (false). On the other hand, a soft clustering approach allows grayscale values to share a degree of membership in several groups. One of the most popular algorithms in the category of crisp clustering is K-means clustering (KMC) algorithm [23,24]. Another example in this category of clustering algorithm is multi-view information-theoretic co-clustering (MV-ITCC), which is based on information-theoretic co-clustering approach [25]. According to recent advances in the development of clustering algorithms, soft clustering algorithms can be divided into fuzzy set [26] based clustering, intuitionistic fuzzy set (IFS) [27] based clustering and neutrosophic set [28] based clustering. One of the most frequently used methods, developed using fuzzy sets, is fuzzy C-means (FCM) [29,30]. Subsequently, researchers introduce several variants of the FCM algorithm. For example, Chen et al. [31] propose multiple-kernel FCM (MKFCM) algorithm using the composite kernel concept. Ji et al. [32] propose the weighted image patch-based FCM (WIPFCM) algorithm, which considers the spatial information of pixels during image clustering. Wang et al. [33] incorporate information-theoretic concept into the FCM algorithm to improve its performance. To deal with the noise in grayscale images, Zhao et al. [34] propose a new version of the FCM algorithm, called generalized fuzzy c-means clustering (GFCC) algorithm. Chaire [35] introduce a novel intuitionistic FCM (IFCM) clustering algorithm by adopting the concept of IFS. Verma et al. [36] propose an improved intuitionistic FCM (IFIFCM) clustering algorithm by including the local spatial information during the clustering. To overcome the drawback of FCM algorithm, Deng et al. [37] propose a new transfer prototype-based fuzzy clustering method. Singh [38] introduces neutrosophic-entropy based clustering algorithm (NEBCA) for performing clustering operation on magnetic resonance imaging (MRI) of Parkinson’s disease.

The above discussed clustering methods [29,31–36,38] are based on the concepts of fuzzy set, IFS and neutrosophic set, and they are able to deal with inherent uncertainties of grayscale values, but they have certain limitations, such as:

(a) The clustering methods based on fuzzy set (i.e., FCM [29] and its variants [31–34]) support the representation of uncertainties of grayscale values only by considering a true degree of membership. These methods define the true degree of membership of a grayscale value in [0, 1].

(b) IFS based clustering methods (i.e., IFCM [35] and IIIFCM [36]) model the uncertain grayscale values with respect to hesitancy. These methods define the hesitancy in terms of true and false degree of memberships.

(c) The clustering method based on neutrosophic sets (i.e., NEBCA [38]) defines the uncertainties of grayscale values with three degrees of memberships, called true, false and indeterministic [39,40].

The above discussion indicates the need of a more robust clustering method that can deal with the drawbacks mentioned in (1)–(4) as well as inherent uncertainties in the CT scan images of COVID-19 patients precisely. Recently, Singh et al. [41] proposed a novel theory to deal with uncertainties, called ambiguous set theory. Ambiguous set theory is better than fuzzy set, IFS and neutrosophic set in terms of its capability of modeling the inherent ambiguities very effectively with four degree of memberships, as true, false, true-ambiguous and false-ambiguous. Singh et al. [41] show the application of this theory in segmenting MRI of human brain. However, the application of this theory has not been extended to the analysis of other types of medical images, such as X-rays and CT scan images. A recent study by Singh and Bose [2] shows that the clustering approach is useful to identify infected regions in CT scan images of COVID-19 patients. With
this motivation and considering the severity of COVID-19, this study extends the application of this theory in the development of a clustering algorithm that could be helpful in the analysis of CT scan images of COVID-19 patients. Hence, the main contributions of this study are five-fold as:

1. First, we first introduce the notion of ambiguous set theory with a mathematical representation.
2. Second, this study presents ambiguous membership functions (AMFs) to define the true, false, true-ambiguous and false-ambiguous memberships of an event using ambiguous set.
3. Third, to quantify the inherent ambiguities associated with the four degree of memberships of ambiguous set (i.e., true, false, true-ambiguous and false-ambiguous), four different entropy formulas are proposed.
4. Fourth, a new image clustering algorithm is proposed in this study using four memberships of the ambiguous set and their respective entropies, called ambiguous D-means fusion clustering algorithm (ADMFCA). The proposed algorithm generates four different clustered images by performing clustering of grayscale values of CT scan images. For clustering, a distance metric, called ambiguous-entropy distance function is introduced, whose main objective is to assign the grayscale values to the different clusters based on the minimum distance criterion. Finally, in order to incorporate all the features of the four different clustered images and obtain a final image, they are aggregated with the help of image fusion [40]. This final image is referred to as final clustered image (FCI). The main reason for using image fusion operation in this algorithm is to integrate the best features into the resultant image [42]. A major problem with most crisp clustering algorithms (as discussed above) is that data points are often assigned to incorrect clusters by ignoring their correlation with some other subset of data points in the problem space [43]. Therefore, the main goal of ADMFCA is to cluster the grayscale values of CT scan images of COVID-19 patients in such a way that correlated features of the infected regions can be easily identified and formed clusters with highly correlated features.
5. Fifth, in support of ambiguous set theory, various definitions, set-theoretical operations, theorems and properties are discussed.

The proposed ADMFCA is validated with chest CT scan images of COVID-19 patients with the corresponding ground truth. The performance of the proposed ADMFCA is compared with five existing clustering algorithms, including KMC [23], FCM [29], GFCM [34], IIFCM [36] and NEBCA [38]. The performance of the proposed ADMFCA and existing algorithms [23,29,34,36,38] is compared using statistical metrics, such as mean squared error (MSE), peak signal-to-noise ratio (PSNR), Dice similarity coefficient (DSC), Jaccard similarity coefficient (JCC) and correlation coefficient (CC). These statistical analyses show the effectiveness of the proposed ADMFCA over the selected clustering algorithms [23,29,34,36,38].

The remainder of this article is organized as follows. Background for the study is presented in Section 2. Section 3 introduces the proposed ambiguous set theory. The proposed ADMFCA for image segmentation is presented in Section 4. Various properties of ambiguous-entropy distance function are discussed in Section 5. Experimental results are described in Section 6. Finally, conclusions and future directions are presented in Section 7.

2. Background for the study

This section presents an overview of the fuzzy set, intuitionistic fuzzy set and neutrosophic set.

2.1. Fuzzy Set

The fuzzy set $\tilde{F}$ for any event $Q(i = 1, 2, \ldots, n)$ in the discrete and finite universe of discourse $S$ can be described as [26]:

$$\tilde{F} = \left\{ \frac{\mu_{\tilde{F}}(Q_1)}{Q_1} + \frac{\mu_{\tilde{F}}(Q_2)}{Q_2} + \ldots + \frac{\mu_{\tilde{F}}(Q_n)}{Q_n} \right\}$$

(3)

For the continuous and infinite universe of discourse $S$, the fuzzy set $\tilde{F}$ for any event $Q(i = 1, 2, \ldots, n)$ can be defined as:

$$\tilde{F} = \left\{ \frac{\mu_{\tilde{F}}(Q_i)}{Q_i} \right\}$$

(4)

In Eqs. (3)-(4), $\mu_{\tilde{F}}(Q_i)$ represents the degree of membership for each $Q_i \in S$. In this theory, the degree of membership of each $Q_i$, i.e., $\mu_{\tilde{F}}(Q_i)$ always belongs to the range $[0, 1]$. In Eqs. (3)-(4), the horizontal bar represents a delimiter. The numerator of each term reflects the degree of membership of each $Q_i$ in the fuzzy set $\tilde{F}$. In Eq. (3), the summation symbol “$+$” represents the aggregation of each $Q_i$, called an aggregation operator. In Eq. (4), the integral sign indicates a continuous function-theoretic aggregation operator for continuous events [26].

2.2. Intuitionistic fuzzy set

Atanassov [27] proposed the concept of intuitionistic fuzzy set (IFS). It helps to represent the hesitancy involved in each $Q_i \in S$ with respect to two membership functions, called degree of membership and degree of non-membership.

For a fixed crisp set $C$, an IFS is denoted as $c^i(Q_i)$, and defined as:

$$c^i(Q_i) = (Q_i, \mu_{c^i}(Q_i), \theta_{c^i}(Q_i))$$

(5)

Here, $\mu_{c^i}(Q_i) \in [0, 1]$ denotes the membership of $Q_i$, and $\theta_{c^i}(Q_i) \in [0, 1]$ denotes the non-membership of $Q_i$ in the $c^i$.

In the IFS, the boundaries of membership and non-membership of $Q_i$ must satisfy the following condition as:

$$0 \leq \mu_{c^i}(Q_i) + \theta_{c^i}(Q_i) \leq 1$$

(6)

The non-membership of $Q_i$ is defined in terms of the fuzzy set as: $\theta_{c^i}(Q_i) = 1 - \mu_{c^i}(Q_i)$. However, this leads to a loss of information when $Q_i$ changes its state from one to another. Therefore, this loss of information is shown using IFS with respect to the membership and non-membership functions as:

$$\Delta(Q_i) = 1 - (\mu_{c^i}(Q_i) + \theta_{c^i}(Q_i))$$

(7)

Here, $\Delta(Q_i)$ indicates the degree of loss. It can be defined only for the IFS, because here $\Delta(Q_i) \neq 0$. However, for ordinary fuzzy set, $\theta_{c^i}(Q_i) = 1 - \mu_{c^i}(Q_i)$, so $\Delta(Q_i) = 0$.

For $Q_i$, an IFS can also be represented in terms of loss as:

$$c^i(Q_i) = (\mu_{c^i}(Q_i), \theta_{c^i}(Q_i), \Delta(Q_i))$$

(8)

2.3. Neutrosophic set

Smarandache [28] introduced the neutrosophic set theory that can model the inherent uncertainty of each $Q_i \in S$ in terms of three degree of memberships, namely truth ($N_T$), indeterminacy ($N_I$) and falsity ($N_F$).

Assume that a neutrosophic set $N$ is defined for the $Q_i$ on the universe of discourse $S$. Here, $N_T$, $N_I$ and $N_F$ for the $Q_i \in S$ can be expressed as: $N_T, N_I, N_F : S \rightarrow [0, 1]^*$. $Q_i = \ldots$
and (False-ambiguous: $\mathbb{E}$) are defined with respect to four degree of membership functions, viz., true, false, true-ambiguous and false-ambiguous, respectively. Since the ambiguities of (True: $\mathbb{E}$) and (False: $\mathbb{E}$) are contradictory, their true and false degree of membership functions are considered complementary to one another. True-ambiguous and false-ambiguous degree of membership functions, on the other hand, are determined by the true and false degree of membership functions, respectively. Thus, the individual values of true, false, true-ambiguous and false-ambiguous, which result from the respective degree of membership functions, indicate the ambiguities of $\mathbb{E}$ in terms of degree-of-true, degree-of-false, degree-of-ambiguity-in-true and degree-of-ambiguity-in-false, respectively. In this way, this theory addresses the problems associated with ambiguous features of the information in the data.

3.2. Proposed ambiguous set

The inherent uncertainty for any event $x$ in the universe of discourse $\mathbb{S}$ can be defined using ambiguous set theory [41], which represents the uncertainty in terms of four degree of membership functions, namely, true ($T$), false ($F$), true-ambiguous ($TA$) and false-ambiguous ($FA$). Mathematically, the ambiguous set can be defined as follows.

**Definition 1 (Ambiguous Set [41]).** An ambiguous set $\mathbb{A}$ on the universe of discourse $\mathbb{S}$ can be defined based on the four membership functions $T, F, TA, FA : \mathbb{S} \rightarrow [0, 1]^+$. These membership functions must satisfy the condition $0 \leq T(x) + F(x) + TA(x) + FA(x) \leq 2^+$ for all $x \in \mathbb{S}$. Here, $T, F, TA$ and $FA$ are real standard values or non-standard subsets of $[0, 1]^+$.

From a philosophical point of view, the ambiguous set takes the value $[0, 1]^+$ on real standard or non-standard subsets. Thus, for engineering applications, instead of taking $[0, 1]^+$, it is useful to take the interval $[0, 1]$, since it is difficult to use $[0, 1]^+$ in real applications such as engineering and science problems.

The AMFs consist of four membership functions, namely, $T, F, TA$ and $FA$. The generalized form of the ambiguous set is called single-valued ambiguous set (SVAS), if the AMFs are singleton subintervals/subsets of the standard real unit interval $[0, 1]$. The SVAS can be defined as:

**Definition 2 (SVAS).** An SVAS $\mathbb{A}$ in $\mathbb{S}$ is represented by four membership functions: $T : \mathbb{S} \rightarrow [0, 1]$, $F : \mathbb{S} \rightarrow [0, 1]$, $TA : \mathbb{S} \rightarrow [0, 1]$ and $FA : \mathbb{S} \rightarrow [0, 1]$. Such SVAS $\mathbb{A}$ can be designated as:

$$
\mathbb{A} = \{T(x), F(x), TA(x), FA(x)|x \in \mathbb{S}\},
$$

with the condition $0 \leq T(x) + F(x) + TA(x) + FA(x) \leq 2^+$, $\forall x \in \mathbb{S}$.

An ambiguous set can be defined for the discrete case as follows.

**Definition 3 (Discrete Ambiguous Set).** An ambiguous set $\mathbb{A}$ for the discrete and finite universe of discourse $\mathbb{S} = \{x_1, x_2, \ldots, x_n\}$ can be represented as:

$$
\mathbb{A} = \left\{ \frac{T(x_1), F(x_1), TA(x_1), FA(x_1)}{x_1}, \frac{T(x_2), F(x_2), TA(x_2), FA(x_2)}{x_2}, \ldots + \frac{T(x_n), F(x_n), TA(x_n), FA(x_n)}{x_n} \right\}
$$

By integrating the four different perceptions of an event, a novel theory is proposed, called ambiguous set theory [41]. According to this theory, the initial perceptions of the event $\mathbb{E}$ are characterized as (True: $\mathbb{E}$) and (False: $\mathbb{E}$). The ambiguities in (True: $\mathbb{E}$) and (False: $\mathbb{E}$) are characterized by the perceptions of (True-ambiguous: $\mathbb{E}$) and (False-ambiguous: $\mathbb{E}$), respectively. The ambiguities in (True: $\mathbb{E}$), (False: $\mathbb{E}$), (True-ambiguous: $\mathbb{E}$) and (False-ambiguous: $\mathbb{E}$) are defined with respect to four degree of membership functions, viz., true, false, true-ambiguous and false-ambiguous, respectively. Since the ambiguities of (True: $\mathbb{E}$) and (False: $\mathbb{E}$) are contradictory, their true and false degree of membership functions are considered complementary to one another. True-ambiguous and false-ambiguous degree of membership functions, on the other hand, are determined by the true and false degree of membership functions, respectively. Thus, the individual values of true, false, true-ambiguous and false-ambiguous, which result from the respective degree of membership functions, indicate the ambiguities of $\mathbb{E}$ in terms of degree-of-true, degree-of-false, degree-of-ambiguity-in-true and degree-of-ambiguity-in-false, respectively. In this way, this theory addresses the problems associated with ambiguous features of the information in the data.

3. The proposed ambiguous set theory

This section presents the philosophy of the ambiguous set, its various definitions followed by related properties.

3.1. Philosophy of ambiguous set

According to the Oxford Dictionary, the word ambiguous is an adjective that means “open to more than one interpretation”. Some information reflects different interpretations, which leads to ambiguity and incompleteness. As a result of this problem, decision-making becomes difficult in most of the cases.

Consider this proposition: “Mr. X is lying”. This statement is either true or false; however, in view of the proposition, the human cognitive process can have the following perceptions:

**Perception 1:** Mr. X is lying.

**Perception 2:** Mr. X is not lying.

Perceptions 1 and 2 can have definite true and false values, respectively. However, there is often uncertainty and incompleteness between truth and falsity. According to human cognitive processes, both uncertainty and incompleteness can be interpreted in terms of the following perceptions:

**Perception 3:** Mr. X is a little lying.

**Perception 4:** It’s a little false that Mr. X is lying.

Perceptions 3 and 4 cannot have distinct true and false values, respectively. Perception 3 is very close to Perception 1, and it inherits the ambiguity from Perception 1, so it can be categorized as having a true-ambiguous value. Similarly, Perception 4 is very close to Perception 2, but it develops ambiguity from Perception 2, so it can be represented with a false-ambiguous value.

The above discussion indicates that any event $\mathbb{E}$ can be viewed in terms of the following four different perceptions listed in P1-P4 as:

**P1:** (True: $\mathbb{E}$) is completely true, i.e., (True: $\mathbb{E}$).

**P2:** (False: $\mathbb{E}$) is completely false, i.e., (False: $\mathbb{E}$).

**P3:** (True: $\mathbb{E}$) is a little true, i.e., (True-ambiguous: $\mathbb{E}$).

**P4:** (False: $\mathbb{E}$) is a little false, i.e., (False-ambiguous: $\mathbb{E}$).

By integrating the four different perceptions of an event, a novel theory is proposed, called ambiguous set theory [41]. According to this theory, the initial perceptions of the event $\mathbb{E}$ are characterized as (True: $\mathbb{E}$) and (False: $\mathbb{E}$). The ambiguities in (True: $\mathbb{E}$) and (False: $\mathbb{E}$) are characterized by the perceptions of (True-ambiguous: $\mathbb{E}$) and (False-ambiguous: $\mathbb{E}$), respectively. The ambiguities in (True: $\mathbb{E}$), (False: $\mathbb{E}$), (True-ambiguous: $\mathbb{E}$) and (False-ambiguous: $\mathbb{E}$)
In Eq. (11), both symbols “−” and “∪” are termed as aggregation operators. For the continuous and infinite $S$, the ambiguous set $A$ can be denoted as:

$$A = \left\{ \left( T(x), F(x), TA(x), FA(x) \right) \right\}$$

(12)

The AMFs can be defined as follows for the SVAS.

**Definition 4 (AMFs).** The four degree of membership functions, namely, $T$, $F$, $TA$ and $FA$ for a SVAS $A$ in $S$ can be mathematically defined as:

$$T(x) = \frac{x - \text{min}(S)}{\text{max}(S) - \text{min}(S)}$$

(13)

$$F(x) = 1 - T(x)$$

(14)

$$TA(x) = \frac{T(x) + FA(x)}{15}$$

(15)

$$FA(x) = \frac{F(x)}{F(x) + FA(x)}$$

(16)

In Eqs. (15) and (16), $A_F$ is termed as the ambiguous distance function. Mathematically, it can be formulated as:

$$A_F(x) = \sqrt{T(x)^2 + F(x)^2}$$

(17)

**Example 1.** A CT scan image consists of grayscale values within the range [0, 255]. Fig. 1(a) and (b) show a CT scan image of a COVID-19 patient and its three different grayscale values at pixel positions $P_{15}$, $P_{17}$ and $P_{112}$, which are 215, 228 and 240, respectively. These three grayscale values create the illusion effect as well as other perceptual problems in terms of their individual intensity. The different grayscale intensities also pose the challenge of distinguishing one region from another. Consequently, users cannot confidently use the linguistic terms “dark grey”, “gray” and “light grey” to describe these three grayscale values. However, this difficulty can be resolved by using ambiguous sets, where inherent imprecision or approximation of grayscale intensities are expressed by AMFs. In this respect, three different ambiguous sets $A_1$, $A_2$ and $A_3$ can be defined for the grayscale values at pixel positions $P_{15}$, $P_{17}$ and $P_{112}$ on the universe of discourse $S = [0, 255]$ using the AMFs (Eqs. (13)-(16)), respectively, as:

$$A_1 = \left\{ \left( T(P_{15}), F(P_{15}), TA(P_{15}), FA(P_{15}) \right) \right\}$$

(18)

$$A_2 = \left\{ \left( T(P_{17}), F(P_{17}), TA(P_{17}), FA(P_{17}) \right) \right\}$$

(19)

$$A_3 = \left\{ \left( T(P_{112}), F(P_{112}), TA(P_{112}), FA(P_{112}) \right) \right\}$$

(20)

In Eq. (18), $T(P_{15})$, $F(P_{15})$, $TA(P_{15})$ and $FA(P_{15})$ indicate the ambiguity belonging to the white pixel, non-white pixel, ambiguous white pixel and ambiguous non-white pixel, respectively. Similar explanations can be provided for Eqs. (19) and (20). Graphical representations of four degree of memberships of $A_1$, $A_2$ and $A_3$ are shown in Fig. 1(c)-(e), respectively.

In Fig. 1(c)-(e), each of the shaded regions is called an ambiguous region (AR). This AR is extremely useful as it clearly describes the inherent ambiguity measured by AMFs. The two-dimensional regions, as shown in Fig. 1(c)-(e), obtain from the presence of ambiguous features of the grayscale values at pixel positions $P_{15}$, $P_{17}$ and $P_{112}$, respectively. Here, the AR provides two valuable information:

1. linguistic description of all uncertainties associated with the effect of AMFs, and
2. the distribution of ambiguity in the two-dimensional plane.

**3.3. Related concepts of ambiguous set**

Entropy can be used to measure the individual ambiguousness represented by the AMFs, namely, $T$, $F$, $TA$ and $FA$. Such measurements of ambiguousness with respect to $T$, $F$, $TA$ and $FA$ are called true entropy (TE), false entropy (FE), true-ambiguous entropy (TAE) and false-ambiguous entropy (FAE), respectively. These four entropies can be defined as follows.

**Definition 5 (Measurements of Ambiguosity).** The four different entropies, viz., TE, FE, TAE and FAE of a SVAS $A$ at $x \in S$ are denoted as a measurement $E_T(A, x)$, $E_F(A, x)$, $E_{TA}(A, x)$ and $E_{FA}(A, x)$, respectively, where $A = \{(x, T(x), F(x), TA(x), FA(x)) | x \in S\}$, which can be defined as follows:

$$E_T(A, x) = -\ln(T(x)) \cdot \ln(T(x))$$

(21)

$$E_F(A, x) = -\ln(F(x)) \cdot \ln(F(x))$$

(22)

$$E_{TA}(A, x) = -\ln(FA(x)) \cdot \ln(FA(x))$$

(23)

$$E_{FA}(A, x) = -\ln(TA(x)) \cdot \ln(TA(x))$$

(24)

**Definition 6 (Operations on Ambiguous Sets).** Let

$$A_1 = \{(x, T_1(x), F_1(x), TA_1(x), FA_1(x)) | x \in S\},$$

$$A_2 = \{(x, T_2(x), F_2(x), TA_2(x), FA_2(x)) | x \in S\},$$

be two ambiguous sets. Some operations on ambiguous sets are given below:

1. $A_1 \subseteq A_2$ if and only if $T_1(x) \leq T_2(x)$, $F_1(x) \geq F_2(x)$, $TA_1(x) \geq TA_2(x)$, and $FA_1(x) \geq FA_2(x)$.

2. $A_1^\dagger = \{(x, T_1(x), F_1(x), TA_1(x), FA_1(x)) | x \in S\},$ where $T_1(x) = F_1(x), F_1(x) = T_1(x), FA_1(x) = 1 - TA_1(x), FA_1(x) = 1 - FA_2(x)$.

3. $A_1 \cap A_2 = \{(x, T_1(x) \land T_2(x)), (F_1(x) \lor F_2(x)), (TA_1(x) \lor TA_2(x)), (FA_1(x) \lor FA_2(x)) | x \in S\}.

4. $A_1 \cup A_2 = \{(x, T_1(x) \lor T_2(x)), (F_1(x) \land F_2(x)), (TA_1(x) \land TA_2(x)), (FA_1(x) \land FA_2(x)) | x \in S\}.

**Definition 7 (Ambiguous Vector and Its Complement),** Let $\vartheta = (\vartheta_1, \vartheta_2, \ldots, \vartheta_n)$ be a vector, where for each $j = 1, 2, \ldots, n$, $\vartheta_j = \{x, T_{\vartheta_0}(x), F_{\vartheta_0}(x), T\vartheta_0(x), TA\vartheta_0(x), FA\vartheta_0(x) | x \in S\}$ is an ambiguous set on the universe $S$. Then, $\vartheta$ is called an ambiguous vector on $S$. We define the complement of $\vartheta$ as $\vartheta^c = (\vartheta_1^c, \vartheta_2^c, \ldots, \vartheta_n^c)$, $\vartheta^c$ denotes the transpose of $\vartheta$. If $n = 1$, we do not distinguish between the ambiguous vector $\vartheta = (\vartheta_1)$ and the ambiguous set $\vartheta_1$.

**Definition 8 (Inner Product),** For each $j = 1, 2, \ldots, n$, let

$$\vartheta_j = \{(x, T_{\vartheta_0}(x), F_{\vartheta_0}(x), T\vartheta_0(x), TA\vartheta_0(x), FA\vartheta_0(x)| x \in S\},$$

$$\theta_j = \{(x, T_{\vartheta_0}(x), F_{\vartheta_0}(x), T\vartheta_0(x), TA\vartheta_0(x), FA\vartheta_0(x)| x \in S\}$$

be ambiguous sets on the universe $S$; let $\vartheta = (\vartheta_1, \vartheta_2, \ldots, \vartheta_n)$, $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$. We call $\vartheta \cdot \theta = \{(x, y, z, u, v) | x \in S\}$ the inner product of $\vartheta$ and $\theta$, where $\vartheta \cdot \theta = \{(x, y, z, u, v) | x \in S\}$, $\theta \cdot \theta = \{(x, y, z, u, v) | x \in S\}$.
Definition 9 (Outer Product). For each \( j = 1, 2, \ldots, n \), let
\[
\Theta_j = \{ (x, T_{\Theta_j}(x), F_{\Theta_j}(x), T_{\Theta_j}(x), T_{\Theta_j}(x), T_{\Theta_j}(x)) | x \in S \},
\]
and
\[
\theta_j = \{ (x, T_{\theta_j}(x), F_{\theta_j}(x), T_{\theta_j}(x), T_{\theta_j}(x), T_{\theta_j}(x)) | x \in S \}.
\]
be ambiguous sets on the universe \( S \); let \( \Theta = (\Theta_1, \Theta_2, \ldots, \Theta_n) \), \( \theta = (\theta_1, \theta_2, \ldots, \theta_n) \). We call \( \Theta \circ \theta = \{ (x, y', z', u', v') | x \in S \} \) the outer product of \( \Theta \) and \( \theta \), where
\[
y' = \bigwedge_{j=1}^{n} (T_{\Theta_j}(x) \lor T_{\theta_j}(x)),
\]
\[
z' = \bigvee_{j=1}^{n} (F_{\Theta_j}(x) \land F_{\theta_j}(x)),
\]
\[
u' = \bigvee_{j=1}^{n} (T_{\Theta_j}(x) \land T_{\theta_j}(x)),
\]
be ambiguous sets on the universe \( S \); let \( \Theta = (\Theta_1, \Theta_2, \ldots, \Theta_n) \), \( \theta = (\theta_1, \theta_2, \ldots, \theta_n) \). Then, \( (\Theta \circ \theta)^c = \Theta^c \circ \theta^c \), \( (\Theta \circ \theta)^c = \Theta^c \circ \theta^c \).

Proof. According to Definitions 6, 8 and 9, we have:

**Theorem 1.** For each \( j = 1, 2, \ldots, n \), let
\[
\Theta_j = \{ (x, T_{\Theta_j}(x), F_{\Theta_j}(x), T_{\Theta_j}(x), T_{\Theta_j}(x), T_{\Theta_j}(x)) | x \in S \},
\]
and
\[
\theta_j = \{ (x, T_{\theta_j}(x), F_{\theta_j}(x), T_{\theta_j}(x), T_{\theta_j}(x), T_{\theta_j}(x)) | x \in S \},
\]
be ambiguous sets on the universe \( S \); let \( \Theta = (\Theta_1, \Theta_2, \ldots, \Theta_n) \), \( \theta = (\theta_1, \theta_2, \ldots, \theta_n) \). Then, \( (\Theta \circ \theta)^c = \Theta^c \circ \theta^c \), \( (\Theta \circ \theta)^c = \Theta^c \circ \theta^c \).
\[(\Theta \cdot \theta)^c = \{ x, \bigvee_{j=1}^n (T_{\Theta_j}(x) \land T_{\theta_j}(x)), \bigwedge_{j=1}^n (F_{\Theta_j}(x) \lor F_{\theta_j}(x)), \\
\land_{j=1}^n (T_{\Theta_j}(x) \lor T_{\theta_j}(x)), \\
\land_{j=1}^n (F_{\Theta_j}(x) \lor F_{\theta_j}(x)) \} | x \in S \}
\]

\[= \{ x, \bigvee_{j=1}^n (T_{\Theta_j}(x) \land T_{\theta_j}(x)), \bigwedge_{j=1}^n (F_{\Theta_j}(x) \lor F_{\theta_j}(x)), \\
1 - \bigwedge_{j=1}^n (T_{\Theta_j}(x) \lor T_{\theta_j}(x)), \\
1 - \bigwedge_{j=1}^n (F_{\Theta_j}(x) \lor F_{\theta_j}(x)) \} | x \in S \}
\]

\[= \Theta^c \circ \theta^c.
\]

\[\Theta \cdot \theta = \left\{ x, \bigvee_{j=1}^n (T_{\Theta_j}(x) \land T_{\theta_j}(x)), \bigwedge_{j=1}^n (F_{\Theta_j}(x) \lor F_{\theta_j}(x)), \\
\land_{j=1}^n (T_{\Theta_j}(x) \lor T_{\theta_j}(x)), \\
\land_{j=1}^n (F_{\Theta_j}(x) \lor F_{\theta_j}(x)) \} | x \in S \}
\]

Proof. According to Definitions 8 and 9:
\[\Theta \cdot \theta = \left\{ x, \bigvee_{j=1}^n (T_{\Theta_j}(x) \land T_{\theta_j}(x)), \bigwedge_{j=1}^n (F_{\Theta_j}(x) \lor F_{\theta_j}(x)), \\
\land_{j=1}^n (T_{\Theta_j}(x) \lor T_{\theta_j}(x)), \\
\land_{j=1}^n (F_{\Theta_j}(x) \lor F_{\theta_j}(x)) \} | x \in S \}
\]

\[= \Theta \circ \theta.
\]

\[\Theta \circ \theta = \left\{ x, \bigvee_{j=1}^n (T_{\Theta_j}(x) \lor T_{\theta_j}(x)), \bigwedge_{j=1}^n (F_{\Theta_j}(x) \land F_{\theta_j}(x)), \\
\land_{j=1}^n (T_{\Theta_j}(x) \lor T_{\theta_j}(x)), \\
\land_{j=1}^n (F_{\Theta_j}(x) \land F_{\theta_j}(x)) \} | x \in S \}
\]

\[= \Theta \circ \theta.
\]

\section*{Definition 10 (Jaccard Similarity Measure)}
Let \(\eta = (\eta_1, \eta_2, \ldots, \eta_n)\) and \(\varsigma = (\varsigma_1, \varsigma_2, \ldots, \varsigma_n)\) be two vectors of length \(n\), where all coordinates are positive. The Jaccard similarity measure of these two vectors is defined as:
\[J(\eta, \varsigma) = \frac{\eta \cdot \varsigma}{\|\eta\|^2 + \|\varsigma\|^2 - \eta \cdot \varsigma}
\]
where \(\eta \cdot \varsigma = \sum_{i=1}^{n} \eta_i \varsigma_i\) is the inner product of the vectors \(\eta\) and \(\varsigma\).

\section*{Definition 11 (Dice Similarity Measure)}
Let \(\eta = (\eta_1, \eta_2, \ldots, \eta_n)\) and \(\varsigma = (\varsigma_1, \varsigma_2, \ldots, \varsigma_n)\) be two vectors of length \(n\), where all coordinates are positive. The Dice similarity measure of these two vectors is defined as:
\[D(\eta, \varsigma) = \frac{2 \eta \cdot \varsigma}{\|\eta\|^2 + \|\varsigma\|^2} = \frac{2 \sum_{i=1}^{n} \eta_i \varsigma_i}{\sum_{i=1}^{n} \eta_i^2 + \sum_{i=1}^{n} \varsigma_i^2}
\]
Definition 12 (Cosine Similarity Measure). Let \( \eta = (\eta_1, \eta_2, \ldots, \eta_n) \) and \( \varsigma = (\varsigma_1, \varsigma_2, \ldots, \varsigma_n) \) be two vectors of length \( n \), where all the coordinates are positive. The cosine similarity measure of these two vectors is defined as:

\[
C(\eta, \varsigma) = \frac{\eta \cdot \varsigma}{\|\eta\| \|\varsigma\|} = \frac{\sum_{i=1}^{n} \eta_i \varsigma_i}{\sqrt{\sum_{i=1}^{n} \eta_i^2} \sqrt{\sum_{i=1}^{n} \varsigma_i^2}}
\]

(27)

Property 1. The Jaccard, Dice and cosine similarity measures satisfy the following properties as:

\[ P_1 : J(\eta, \varsigma), D(\eta, \varsigma), C(\eta, \varsigma) \in [0, 1] \]

\[ P_2 : J(\eta, \varsigma) = J(\varsigma, \eta), D(\eta, \varsigma) = D(\varsigma, \eta), C(\eta, \varsigma) = C(\varsigma, \eta) \]

\[ P_3 : J(\eta, \varsigma) = \frac{D(\eta, \varsigma)}{\eta \cup \varsigma} = \frac{\eta \cap \varsigma}{\eta \cup \varsigma} = 1 \]

The above similarity measures motivate the following definition.

Definition 13. Let

\[ \mathcal{A}_1 = \{x, T_1(x), F_1(x), TA_1(x), FA_1(x)\} | x \in \mathbb{S} \]

\[ \mathcal{A}_2 = \{x, T_2(x), F_2(x), TA_2(x), FA_2(x)\} | x \in \mathbb{S} \]

be two ambiguous sets on the universe \( \mathbb{S} \). Let \( \eta = (\eta_1, \eta_2, \ldots, \eta_n) \) and \( \varsigma = (\varsigma_1, \varsigma_2, \ldots, \varsigma_n) \) be two vectors of length \( n \), where \( \eta_i, \varsigma_i \in \mathbb{S} \) for \( i = 1, 2, \ldots, n \). Let \( w_1, w_2, \ldots, w_n \) be non-negative real numbers, called weights. The ambiguous weighted jaccard similarity measure, ambiguous weighted Dice similarity measure and ambiguous weighted cosine similarity measure of these two ambiguous sets for the vectors \( \eta, \varsigma \) are defined, respectively, as:

\[
AWJ(\mathcal{A}_1, \mathcal{A}_2; \eta, \varsigma) = \sum_{i=1}^{n} w_i J(T_1(x_i), F_1(x_i), TA_1(x_i), FA_1(x_i)),
\]

\[
AWD(\mathcal{A}_1, \mathcal{A}_2; \eta, \varsigma) = \sum_{i=1}^{n} w_i D(T_1(x_i), F_1(x_i), TA_1(x_i), FA_1(x_i)),
\]

\[
AWC(\mathcal{A}_1, \mathcal{A}_2; \eta, \varsigma) = \sum_{i=1}^{n} w_i C(T_1(x_i), F_1(x_i), TA_1(x_i), FA_1(x_i)).
\]

4. The proposed ADMFCA

This section introduces the proposed ADMFCA for clustering grayscale images. The proposed ADMFCA is based on ambiguous set theory, entropies and image fusion operation. Each step of the proposed ADMFCA is explained next.

Step 1. Define the grayscale domain of image: The grayscale value \( G \) associated with each pixel \( P_G(i = 1, 2, \ldots, m | j = 1, 2, \ldots, n) \) of an input gray image \( I_G \) can be expressed in a grayscale domain as:

\[
I_G = \begin{bmatrix}
G_{11} & G_{12} & \ldots & G_{1n} \\
G_{21} & G_{22} & \ldots & G_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
G_{m1} & G_{m2} & \ldots & G_{mn}
\end{bmatrix}
\]

(28)

Step 2. Define the ambiguous domain of image: The ambiguous domain for the grayscale image \( I_G \) is defined by representing the grayscale value \( G \) of each pixel in the ambiguous set. The ambiguous set of each \( G \) is denoted as \( A_G \), and can be expressed in the following matrix \( A_G \) as:

\[
\begin{bmatrix}
A_{G11} & A_{G12} & \ldots & A_{G1n} \\
A_{G21} & A_{G22} & \ldots & A_{G2n} \\
\vdots & \vdots & \ddots & \vdots \\
A_{Gm1} & A_{Gm2} & \ldots & A_{Gmn}
\end{bmatrix}
\]

(29)

In Eq. (29), each \( A_G \) is defined as:

\[
A_G = \{G_{ij}, (T(G_{ij}), F(G_{ij}), TA(G_{ij}), FA(G_{ij})) | G_{ij} \in \mathbb{S}\}
\]

(30)

In Eq. (30), the four AMFs, namely, \( T, F, TA \) and \( FA \) for \( G_{ij} \) in \( \mathbb{S} \) can be defined as:

\[
T(G_{ij}) = G_{ij} - \min(\mathbb{S})
\]

\[
F(G_{ij}) = 1 - T(G_{ij})
\]

\[
TA(G_{ij}) = \frac{T(G_{ij})}{\max(\mathbb{S}) - \min(\mathbb{S})}
\]

\[
FA(G_{ij}) = \frac{F(G_{ij})}{\max(\mathbb{S}) - \min(\mathbb{S})}
\]

(31) (32) (33) (34)

In Eq. (31), \( \min \) and \( \max \) represent the minimum and maximum functions, respectively. In Eqs. (33) and (34), the ambiguous distance function \( AD \) can be defined as:

\[
AD(G_{ij}) = \sqrt{T(G_{ij})^2 + F(G_{ij})^2}
\]

(35)

Step 3. Measurements of ambiguousness for ambiguous set: The ambiguousness of the AMFs (Eqs. (31)–(34)) can be measured individually by \( TE, TF, TAE \) and \( FAE \), which are expressed as \( E_T(A_G, G_{ij}), E_F(A_G, G_{ij}), E_{TA}(A_G, G_{ij}) \) and \( E_{FA}(A_G, G_{ij}) \), defined in Eqs. (36)–(39), respectively, as:

\[
E_T(A_G, G_{ij}) = -T(G_{ij}) \cdot \ln(T(G_{ij}))
\]

\[
E_F(A_G, G_{ij}) = -F(G_{ij}) \cdot \ln(F(G_{ij}))
\]

\[
E_{TA}(A_G, G_{ij}) = -TA(G_{ij}) \cdot \ln(TA(G_{ij}))
\]

\[
E_{FA}(A_G, G_{ij}) = -FA(G_{ij}) \cdot \ln(FA(G_{ij}))
\]

(36) (37) (38) (39)

Step 4. Selection of clusters for the entropies: Choose \( D \) initial number of clusters at random for the \( E_T(A_G, G_{ij}), E_F(A_G, G_{ij}), E_{TA}(A_G, G_{ij}) \) and \( E_{FA}(A_G, G_{ij}) \) (Eqs. (36)–(39)), respectively, as \( C^{T}_{d}, C^{F}_{d}, C^{TA}_{d} \) and \( C^{FA}_{d} \), respectively; where \( d = 1, 2, \ldots, D \). Here, \( C^{T}_{d}, C^{F}_{d}, C^{TA}_{d} \) and \( C^{FA}_{d} \) occupy each cluster \( E_T(A_G), E_F(A_G), E_{TA}(A_G) \) and \( E_{FA}(A_G) \), respectively.

Step 5. Define the set of centers for each of the clusters: Define a set of random initialized centers for each of the clusters \( C^{T}_{d}, C^{F}_{d}, C^{TA}_{d} \) and \( C^{FA}_{d} \) as:

\[
W(0) = [W_1(0), W_2(0), \ldots, W_D(0)] \in C^{T}_{d}
\]

\[
X(0) = [X_1(0), X_2(0), \ldots, X_D(0)] \in C^{F}_{d}
\]

\[
Y(0) = [Y_1(0), Y_2(0), \ldots, Y_D(0)] \in C^{TA}_{d}
\]

\[
Z(0) = [Z_1(0), Z_2(0), \ldots, Z_D(0)] \in C^{FA}_{d}
\]

(40) (41) (42) (43)

Here, 0 indicates the 1st epoch of the algorithm. From Eq. (40), it can be assumed that \( C^{T}_{d} \) cluster has \( W_1 \) center, where \( W_i \leq E_T(A_G, G_{ij}) \), \( 1 \leq d \leq D \). Similar assumptions can be made for \( C^{F}_{d}, C^{TA}_{d} \) and \( C^{FA}_{d} \) in terms of Eqs. (41)–(43), respectively.

Step 6. Set the epochs: For individual clustering of \( E_T(A_G, G_{ij}), E_F(A_G, G_{ij}), E_{TA}(A_G, G_{ij}) \) and \( E_{FA}(A_G, G_{ij}) \), the epoch \( e \) from 0 to \( \text{Epoch} \) is set as \( e = 0, 1, \ldots, \text{Epoch} \), where \( \text{Epoch} \) denotes the maximum number of epochs.
Step 7. Computation of distances between entropies and centers: Each of the entropies \( E_T(A_1, G_j) \), \( E_T(A_2, G_j) \), \( E_T(A_3, G_j) \) and \( E_T(A_4, G_j) \) is assigned to the individual clusters \( C^T_1 \), \( C^T_2 \), \( C^T_3 \) and \( C^T_4 \) with respect to the nearest clusters \( W_i(0) \), \( X_i(0) \), \( Y_i(0) \) and \( Z_i(0) \), respectively. The determination of the nearest center vectors \( W_i(0), X_i(0), Y_i(0) \) and \( Z_i(0) \) is done by employing ambiguous-entropy distance function. The proposed function computes the distance between \( E_T(A_j, G_j) \) and \( W_i(0) \) as:

\[
\text{Dist}[E_T(A_j, G_j), W_i(0)] = \left[ E_T(A_j, G_j)^2 + |W_i(0)|^2 - 2 \cdot [E_T(A_j, G_j)] \cdot |W_i(0)| \right]
\]

(44)

Similarly, the proposed metric computes the distances between \( E_T(A_j, G_j) \) and \( X_i(0) \), \( E_T(A_j, G_j) \) and \( Y_i(0) \), and \( E_T(A_j, G_j) \) and \( Z_i(0) \), defined in Eqs. (45)–(47), respectively, as:

\[
\begin{align*}
\text{Dist}[E_T(A_j, G_j), X_i(0)] & = \left[ E_T(A_j, G_j)^2 + |X_i(0)|^2 - 2 \cdot [E_T(A_j, G_j)] \cdot |X_i(0)| \right] \\
\text{Dist}[E_T(A_j, G_j), Y_i(0)] & = \left[ E_T(A_j, G_j)^2 + |Y_i(0)|^2 - 2 \cdot [E_T(A_j, G_j)] \cdot |Y_i(0)| \right] \\
\text{Dist}[E_T(A_j, G_j), Z_i(0)] & = \left[ E_T(A_j, G_j)^2 + |Z_i(0)|^2 - 2 \cdot [E_T(A_j, G_j)] \cdot |Z_i(0)| \right]
\end{align*}
\]

(45)–(47)

In Eqs. (44)–(47), \( \text{Dist} \) denotes the ambiguous-entropy distance metric. In Eq. (44), if \( W_i(0) \) is the closest center to \( E_T(A_j, G_j) \), then it is assigned to the cluster \( C^T_1 \). A similar explanation can be given for Eqs. (45)–(47).

Step 8. Selection criterion of clusters: The selection of each cluster by \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), and \( E_T(A_j, G_j) \) depends on the minimum values of the ambiguous-entropy distances (Eqs. (44)–(47), respectively). For example, let \( W_i(0) \) and \( W_j(0) \) be the two randomly defined centers for the clusters \( C^T_1 \) and \( C^T_2 \) with respect to the clustering \( E_T(A_j, G_j) \). Now, \( E_T(A_j, G_j) \in W_i(0) \) if it satisfies the following condition as:

\[
\text{Dist}[E_T(A_j, G_j), W_i(0)] < \text{Dist}[E_T(A_j, G_j), W_j(0)]
\]

(48)

where, \( W_i(0) \neq W_j(0) \). In Eq. (48), \( \text{Dist}[E_T(A_j, G_j), W_i(0)] \) and \( \text{Dist}[E_T(A_j, G_j), W_j(0)] \) can be obtained by employing Eq. (44). Eq. (48) indicates that \( W_i(0) \) is the nearest center for \( E_T(A_j, G_j) \), so it is assigned to the cluster \( C^T_1 \). A similar explanation can be given for clustering \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), and \( E_T(A_j, G_j) \), the respective new centers are denoted as \( W_i(e+1), X_i(e+1), Y_i(e+1), Z_i(e+1) \), defined in Eqs. (49)–(52), respectively, as:

\[
W_i(e+1) = \frac{1}{A \times B} \sum_{i=1}^{A} \sum_{j=1}^{B} E_T(A_j, G_j)
\]

(49)

\[
X_i(e+1) = \frac{1}{I \times J} \sum_{i=1}^{I} \sum_{j=1}^{J} E_T(A_j, G_j)
\]

(50)

\[
Y_i(e+1) = \frac{1}{P \times Q} \sum_{i=1}^{P} \sum_{j=1}^{Q} E_T(A_j, G_j)
\]

(51)

\[
Z_i(e+1) = \frac{1}{S \times T} \sum_{i=1}^{S} \sum_{j=1}^{T} E_T(A_j, G_j)
\]

(52)

In Eqs. (49)–(52), \( A \times B, I \times J, P \times Q \) and \( S \times T \) represent the size of the clusters \( C^T_1 \), \( C^T_2 \), \( C^T_3 \) and \( C^T_4 \), respectively.

Step 10. Stop the clustering process: Go to Step 6 and proceed from epoch \( e = 0 \) to the next epoch \( e = e + 1 \). This process continues until the centers stop changing or the algorithm reaches the maximum epoch \( E_p \).

Step 11. Generate the clustered images: Individual clustering of \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \) and \( E_T(A_j, G_j) \) generates the four different clustered images, called TE clustered image (TECI), FE clustered image (FECI), TAE clustered image (TAEICI) and FAE clustered image (FAEICI). TECI, FECI, TAEICI and FAEICI are denoted as \( ECI \), \( ECI \), \( ECI \) and \( ECI \), respectively.

Step 12. Obtain the final clustered image: The final clustered image (FCI) is generated by applying the image fusion operation [40] on four clustered images, viz., \( ECI \), \( ECI \), \( ECI \) and \( ECI \) as:

\[
FCI = \frac{1}{4} [TECI + FECI + TAEICI + FAEICI]
\]

(53)

Here, \( FCI \) denotes the FCI.

The pseudocode of the proposed ADMFCA is summarized in Algorithm 1.

Algorithm 1 PROCEDURE ADMFCA().

Input: an image \( C_i \) with grayscale value \( G_{ij} \in P_i \) is defined within the range \( S = [0, G] \) with \( G = 255 \), where each \( P_i(l = 1, 2, \ldots, m) \) denotes the pixel of \( C_i \).

Output: final clustered image (FCI).

1: Represent \( C_i \) by grayscale domain denoted by \( GCI \) (Eq. (28)).

2: Represent \( C_i \) into ambiguous domain denoted by \( A_G \) (Eq. (29)).

3: Measure the ambiguousness of the AMFs of \( A_G \), and expressed as \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \) and \( E_T(A_j, G_j) \) (Eqs. (36)–(39), respectively).

4: Choose \( D \) initial clusters at random for the \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \), \( E_T(A_j, G_j) \) and \( E_T(A_j, G_j) \) as \( C^T_1, C^T_2, C^T_3 \) and \( C^T_4 \), respectively; where \( d = 1, 2, \ldots, D \).

5: Define a set of random initialized centers for each of the clusters \( C^T_1, C^T_2, C^T_3 \) and \( C^T_4 \) (Eqs. (40)–(43), respectively), while \( e = 0 \) do

a: Compute the distances between entropies and centers (Eqs. (44)–(47)).

b: Select each cluster by employing condition given (Eq. (48)).

c: Update each of the centers (Eqs. (49)–(52)).

end

6: Generate the four different clustered images as \( TECI \), \( FECI \), \( TAEICI \) and \( FAEICI \).

7: Apply the image fusion operation on four clustered images to obtain the final clustered image as \( FCI \).

5. Properties of ambiguous-entropy distance function

This section presents various properties of ambiguous-entropy distance function. This function is used in the proposed ADMFCA to compute the distance between entropy and center of the clusters. Consider the following generalized form of ambiguous-entropy distance function that computes the distance between \( E_T(A_j, G_j) \) and \( W_i \) as:

\[
\text{Dist}[E_T(A_j, G_j), W_i] = \left[ E_T(A_j, G_j)^2 + |W_i|^2 - 2 \cdot [E_T(A_j, G_j)] \cdot |W_i| \right]
\]

(54)

Here, ambiguous-entropy distance function \( \text{Dist}[E_T(A_j, G_j), W_i] \) is used to compute the distance between \( E_T(A_j, G_j) \) and \( W_i \). For ease of explanation of various properties of this function, we only
consider the vectors $E_i(Å_j, G_j)$ and $W_i$. However, these properties are also valid in the case of the computation of the distances between $E_i(Å_j, G_j)$ and $X_i$, $E_{i\mathbf{a}}(Å_j, G_j)$ and $Y_i$, and $E_{i\mathbf{a}}(Å_j, G_j)$ and $Z_i$. In the following, we have discussed various properties of ambiguous-entropy distance function in terms of Eq. (54).

**Property 2.** $\text{Dist}[E_i(Å_j, G_j), W_i] : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is equal to 0 if $E_i(Å_j, G_j) = W_i$.

**Proof.** If $E_i(Å_j, G_j) = W_i$ then $\text{Dist}[E_i(Å_j, G_j), W_i] = \|E_i(Å_j, G_j) - W_i\|^2 = 0$.

**Property 3.** As the distance between $E_i(Å_j, G_j)$ and $W_i$ increases, $\text{Dist}[E_i(Å_j, G_j), W_i] > 0$.

**Proof.** Assume two centers $W_i$ and $W_j$, where $W_i$, $W_j \in W$ and $W_i > W_j$. From Eq. (54), it is clear that $\text{Dist}[E_i(Å_j, G_j), W_j] > \text{Dist}[E_i(Å_j, G_j), W_i]$. It indicates that as $W_i$ increases, $\text{Dist}[E_i(Å_j, G_j), W_i]$ increases.

Similarly, assume two TEs as $E_i(Å_j, G_j)$ and $E_i(Å_{ij}, G_{ij})$, where $E_i(Å_j, G_j) > E_i(Å_{ij}, G_{ij})$. Now, if we compute the distances of these two TEs with respect to $W_i$, then clearly get $\text{Dist}[E_i(Å_j, G_j), W_i] > \text{Dist}[E_i(Å_{ij}, G_{ij}), W_i]$. It indicates that as $E_i(Å_j, G_j)$ increases, $\text{Dist}[E_i(Å_j, G_j), W_i]$ increases.

**Property 4.** The ambiguous-entropy distance function follows the symmetry property, i.e., $\text{Dist}[E_i(Å_j, G_j), W_i] = \text{Dist}[W_i, E_i(Å_j, G_j)]$.

**Proof.** It is obvious from Eq. (54).

**Theorem 3.** For $\text{Dist}[E_i(Å_j, G_j), W_i] : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, there exists an Euclidean space $E$ and a distance vector $V : \mathbb{R} \times \mathbb{R} \rightarrow E$ such that $\text{Dist}[E_i(Å_j, G_j), W_i] = \|V(E_i(Å_j, G_j), W_i)\|$, where $\|\cdot\|$ denotes the dot product of two vectors.

**Proof.** Let $E$ be the 1-dimensional Euclidean space $E$. Define $V : \mathbb{R} \times \mathbb{R} \rightarrow E$ by $V(a, b) = a - b$. Then,

\[
\begin{align*}
\text{Dist}[E_i(Å_j, G_j), W_i] &= \|V(E_i(Å_j, G_j), W_i)\|^2, \\
&= \|V(E_i(Å_j, G_j) - W_i)\|^2, \\
&= \|E_i(Å_j, G_j) - W_i\|^2 = \text{Dist}[E_i(Å_j, G_j), W_i].
\end{align*}
\]

**Theorem 4.** $\text{Dist}[E_i(Å_j, G_j), W_i] = \|V(E_i(Å_j, G_j), W_i)\|^2$, where $\|\cdot\|$ denotes the standard Euclidean norm on $\mathbb{R}^2$.

**Proof.** By the previous theorem, we have

\[
\text{Dist}[E_i(Å_j, G_j), W_i] = \|V(E_i(Å_j, G_j), W_i)\|^2 = \|V(E_i(Å_j, G_j) - W_i)\|^2 = \text{Dist}[E_i(Å_j, G_j), W_i].
\]

6. Experimental results

This section is divided into several subsections, such as: dataset description is provided in Section 6.1. Various performance evaluation metrics are discussed in Section 6.2. The initial experimental setup is presented in Section 6.3. Visual analysis of the clustered images is discussed in Section 6.4. Finally, the comparison with existing models is presented in Section 6.5.

### Table 2

| Group | Label | Extracted CT scan image | Image size in KB (before preprocessed) | Image size in KB (after preprocessed) |
|-------|-------|--------------------------|----------------------------------------|--------------------------------------|
| 1     | #142  | 169                      | 129                                    |
| 2     | #94   | 237                      | 172                                    |
| 3     | #105  | 218                      | 154                                    |
| 4     | #85   | 202                      | 146                                    |
| Group #1 | 5     | #100                     | 207                                    |
|       | 6     | #110                     | 211                                    |
|       | 7     | #94                      | 212                                    |
|       | 8     | #96                      | 184                                    |
|       | 9     | #109                     | 181                                    |
|       | 10    | #155                     | 173                                    |
| Group #2 | 1     | #118                     | 179                                    |
|       | 2     | #106                     | 231                                    |
|       | 3     | #81                      | 216                                    |
|       | 4     | #71                      | 191                                    |
|       | 5     | #76                      | 193                                    |
|       | 6     | #87                      | 211                                    |
|       | 7     | #113                     | 225                                    |
|       | 8     | #120                     | 169                                    |
|       | 9     | #90                      | 173                                    |
|       | 10    | #179                     | 177                                    |
| Group #3 | 1     | #129                     | 188                                    |
|       | 2     | #97                      | 250                                    |
|       | 3     | #92                      | 223                                    |
|       | 4     | #95                      | 206                                    |
|       | 5     | #87                      | 210                                    |
|       | 6     | #98                      | 222                                    |
|       | 7     | #102                     | 222                                    |
|       | 8     | #107                     | 192                                    |
|       | 9     | #100                     | 185                                    |
|       | 10    | #166                     | 179                                    |
| Group #4 | 1     | #136                     | 173                                    |
|       | 2     | #86                      | 223                                    |
|       | 3     | #100                     | 190                                    |
|       | 4     | #77                      | 179                                    |
|       | 5     | #80                      | 179                                    |
|       | 6     | #92                      | 191                                    |
|       | 7     | #89                      | 177                                    |
|       | 8     | #114                     | 181                                    |
|       | 9     | #104                     | 185                                    |
|       | 10    | #171                     | 163                                    |

6.1. Dataset description

The proposed ADMFCA and selected clustering algorithms, namely KMC [23], FCM [29], GFCM [34], IIFCM [36] and NEBCA [38] are applied to different types of chest CT scan images of COVID-19 patients [45]. This dataset contains CT scan images of COVID-19 patients with 20 different labels. In this study, CT scan images with 10 different labels are selected for the experiment. Out of each label, four different CT scan images are selected. Thus, a total of $10 \times 4 = 40$ CT scan images are available with their respective ground truths. These 40 images are split into four different groups, called Group #1, Group #2, Group #3 and Group #4. However, the extracted CT scan images have noise and poor resolution issues. Therefore, these images are preprocessed before carrying out the experiment. The adaptive filtering technique [46] and the histogram equalization method [47] are used for noise removal and resolution improvement, respectively. Eventually, these preprocessed images are used for the experiment. Detailed information on the experimental datasets is available in Table 2.

6.2. Performance evaluation metrics

Clustered images obtained using the proposed ADMFCA and selected clustering algorithms are evaluated using five statistical metrics, namely MSE, PSNR, DSC, JSC and CC. These metrics are
defined based on the input grayscale image \((I_G)\), final clustered image \((F_C)\) and corresponding ground truth \((G_T)\) as:

\[
\text{MSE}(I_G, F_C) = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} (I_{Gm} - F_{Cm})^2
\]

\[
\text{PSNR}(I_G, F_C) = 10 \times \log_{10} \left( \frac{255^2}{\text{MSE}(I_G, F_C)} \right)
\]

\[
\text{DSC}(F_C, G_T) = \frac{2|F_C \cap G_T|}{|F_C + G_T|}
\]

\[
\text{JSC}(F_C, G_T) = \frac{|F_C \cap G_T|}{|F_C|}
\]

\[
\text{CC}(F_C, G_T) = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} (F_{Cm} - \bar{F}_C)(G_{Tn} - \bar{G}_T)}{\sqrt{\sum_{m=1}^{M} \sum_{n=1}^{N} (F_{Cm} - \bar{F}_C)^2 \sum_{m=1}^{M} \sum_{n=1}^{N} (G_{Tn} - \bar{G}_T)^2}}
\]

A lower MSE value (Eq. (55)) indicates a lower intensity loss and produces a robust clustered image. A higher PSNR value (Eq. (56)) shows less distortion in grayscale values and generates a better clustered image. The DSC and JSC values (Eqs. (57) and (58), respectively) always lie in the range of 0–1. Their values close to 1 imply that the region of interest of the clustered image is nearly similar to the corresponding ground truth. The CC value (Eq. (59)) remains within the range \([-1, 1]\). In Eq. (59), \(F_C\) and \(G_T\) indicate mean values of \(F_C\) and \(G_T\), respectively. A CC value close to 1 shows that the region of interest of the clustered image is perfectly similar to its corresponding background truth.

6.3. Initial experimental set-up

The main objective of ADMFCA is to effectively detect infected regions by clustering the grayscale values of the CT scan images of COVID-19 patients. In this algorithm, we first define the grayscale values \((G_T)\) of all preprocessed CT scan images of COVID-19 patients (Table 2) in terms of the four degree of memberships in the ambiguous set as:

- \(G_{ij}\) of all white pixels are defined by \(T(G_{ij})\),
- \(G_{ij}\) of all non-white pixels are defined by \(F(G_{ij})\),
- \(G_{ij}\) of all white pixels with certain non-white pixels are defined by \(T(A(G_{ij}))\), and
- \(G_{ij}\) of all non-white pixels with certain white pixels are represented by \(F(A(G_{ij}))\).

The individual ambiguity of \(T(G_{ij})\), \(F(G_{ij})\), \(T(A(G_{ij}))\) and \(F(A(G_{ij}))\) is measured using the corresponding TE, FE, TAE and FAE (Eqs. (36)–(39), respectively). Finally, the proposed algorithm is applied sequentially to create clusters \(T(G_{ij})\), \(F(G_{ij})\), \(T(A(G_{ij}))\), and \(F(A(G_{ij}))\). The proposed algorithm develops four different clustered images, called TECI, FECI, TAEI and FAEI. The final clustered images (FCIs) are obtained by applying the image fusion operation (Eq. (53)) to TECI, FECI, TAEI and FAEI.

The proposed algorithm is simulated by selecting three different cluster numbers as \(D = 2, 3, 4\). The main objective of the simulation with different cluster numbers is to determine which cluster number is best to generate the optimal FCIs. The best cluster number for the proposed ADMFCA is determined by
Fig. 3. Visual analysis of clustering the CT scan image (Group: #1, Label: 1, Image: #142) based on the ADMFCA: (a) original CT scan image, (b) ground truth of (a), (c) preprocessed CT scan image of (a), (d) TECL, (e) FECL, (f) TAECL, (g) FAECL, (h) FCL, (i) histogram of (c), and (j) histogram of (h).
FCIs are obtained by setting the maximum number of epochs to \( M \) as MSE, PSNR, DSC, JSC, and CC (Eqs. (55)–(59), respectively). The clustering operation on grayscale values is applied to the preprocessed image (Fig. 2(c)) to obtain the FCIs with respect to \( D = 3 \). Finally, the MSE, PSNR, DSC, and CC values of the FCIs are significantly better than the FCIs (Fig. 2(d) and (f)) obtained by ADMFCA. The MSE, PSNR, DSC, JSC, and CC values in the pixels before performing the clustering operation.

6.4. Visual analysis

A visual analysis is conducted to assess the quality of FCIs obtained from preprocessed CT scan images of COVID-19 patients. To demonstrate visual analysis, CT scan images #142 (Label: 1), #118 (Label: 1), #166 (Label: 10) and #114 (Label: 8) are selected from Groups #1–#4, respectively. These selected images are shown in Figs. 3–6(a). The ground truths of the selected images are shown in Figs. 3–6(b). The four different clustered images, namely TECIs, FECIs, TAECIs, and FAECIs, obtained by the proposed ADMFCA are depicted in Figs. 3–6(d)–(g), respectively. By applying the fusion operation to TECI, FECI, TAECI, and FAECI, we obtain the FCIs shown in Figs. 3–6(h). In Figs. 3–6(h), dense white pixels in the lungs denote infection and symptoms of COVID-19. Figs. 3–6(i) present the histograms of the preprocessed images (Figs. 3–6(c)), which indicate the distribution of grayscale values in the pixels before performing the clustering operation. Figs. 3–6(j) depict the histograms of the FCIs (Figs. 3–6(h)), which indicate the distribution of grayscale values in the pixels after performing the clustering operation. From Figs. 3–6(j), it can be seen that the proposed ADMFCA has effectively performed the clustering operation on grayscale values.

Table 3

| Group  | Label  | CT scan image (preprocessed) | KMC | FCM | GFCM | IIFCM | NEBCA | ADMFCA |
|--------|--------|-----------------------------|-----|-----|------|-------|-------|--------|
|        |        |                             | 1.01| 1.02| 1.03 | 1.04  | 1.05  | 1.06   |
| Group #1 | #118 | 497.78                       | 313.26 | 213.13 | 135.13 | 114.13 | 107.13 | 1.04   |
|        | #106 | 497.64                       | 312.25 | 216.40 | 127.40 | 116.14 | 102.14 | 1.02   |
|        | #81  | 425.53                       | 314.63 | 263.64 | 136.64 | 114.64 | 103.40 | 1.03   |
|        | #71  | 470.01                       | 316.22 | 239.11 | 136.11 | 114.11 | 101.11 | 1.04   |
|        | #76  | 594.30                       | 317.85 | 212.13 | 123.13 | 112.13 | 103.20 | 1.03   |
|        | #87  | 593.02                       | 314.75 | 236.87 | 125.87 | 117.27 | 217.27 | 2.01   |
|        | #113 | 438.52                       | 315.33 | 235.77 | 128.77 | 115.77 | 104.77 | 1.03   |
|        | #120 | 450.75                       | 317.63 | 238.28 | 126.28 | 115.28 | 104.28 | 1.04   |
|        | #90  | 435.72                       | 312.82 | 232.82 | 132.82 | 112.82 | 105.82 | 1.05   |
|        | #179 | 493.02                       | 314.75 | 236.87 | 126.87 | 116.87 | 104.87 | 1.04   |
| Group #2 | #129 | 416.24                       | 311.10 | 239.67 | 135.17 | 119.54 | 104.04 | 1.04   |
|        | #97  | 430.72                       | 312.13 | 230.24 | 121.24 | 110.15 | 105.15 | 1.05   |
|        | #92  | 563.17                       | 302.58 | 238.11 | 128.11 | 113.21 | 106.21 | 1.06   |
|        | #95  | 583.16                       | 316.24 | 215.31 | 124.31 | 115.11 | 104.11 | 1.04   |
|        | #87  | 511.18                       | 339.95 | 212.42 | 129.42 | 114.23 | 105.23 | 1.05   |
|        | #98  | 475.31                       | 312.15 | 266.57 | 129.23 | 116.45 | 103.04 | 1.03   |
|        | #102 | 411.15                       | 311.16 | 218.77 | 127.15 | 118.12 | 218.12 | 2.05   |
|        | #107 | 415.15                       | 324.17 | 240.32 | 131.31 | 111.13 | 301.30 | 3.01   |
|        | #100 | 435.13                       | 321.91 | 238.83 | 138.93 | 118.93 | 104.93 | 1.04   |
|        | #166 | 413.17                       | 321.12 | 244.87 | 136.75 | 117.75 | 102.75 | 2.02   |
| Group #3 | #136 | 425.64                       | 321.20 | 239.67 | 129.12 | 119.14 | 204.14 | 1.04   |
|        | #86  | 411.72                       | 332.13 | 231.14 | 125.21 | 110.25 | 203.25 | 1.03   |
|        | #100 | 464.17                       | 312.38 | 248.21 | 128.21 | 113.23 | 405.23 | 4.05   |
|        | #77  | 583.16                       | 331.14 | 236.32 | 139.11 | 115.21 | 102.21 | 1.02   |
|        | #80  | 511.38                       | 341.95 | 232.45 | 128.12 | 114.24 | 102.24 | 1.02   |
|        | #92  | 477.36                       | 322.53 | 236.67 | 133.43 | 116.41 | 106.31 | 1.06   |
|        | #89  | 414.35                       | 324.66 | 218.71 | 139.25 | 118.22 | 102.22 | 1.02   |
|        | #114 | 418.35                       | 314.27 | 239.31 | 139.32 | 112.23 | 205.23 | 2.05   |
|        | #104 | 436.23                       | 321.92 | 236.82 | 130.82 | 118.92 | 103.03 | 1.03   |
|        | #171 | 413.27                       | 323.23 | 214.82 | 136.54 | 114.17 | 205.17 | 2.05   |
| Group #4 | #138 | 464.38                       | 324.11 | 232.91 | 133.42 | 115.50 | 159.50 | 1.59   |

Evaluating the quality of the FCIs using statistical metrics, such as MSE, PSNR, DSC, JSC and CC (Eqs. (55)–(59), respectively). The FCIs are obtained by setting the maximum number of epochs to \( \text{Epoch} = 100 \).
Fig. 4. Visual analysis of clustering the CT scan image (Group: #2, Label: 1, Image: #118) based on the ADMFCA: (a) original CT scan image, (b) ground truth of (a), (c) preprocessed CT scan image of (a), (d) TECL, (e) FECL, (f) TAECI, (g) FAECI, (h) FCI, (i) histogram of (c), and (j) histogram of (h).
Fig. 5. Visual analysis of clustering the CT scan image (Group: #3, Label: 10, Image: #166) based on the ADMFCA: (a) original CT scan image, (b) ground truth of (a), (c) preprocessed CT scan image of (a), (d) TECl, (e) FECl, (f) TAECI, (g) FAECI, (h) FCI, (i) histogram of (c), and (j) histogram of (h).
Fig. 6. Visual analysis of clustering the CT scan image (Group: #4, Label: 8, Image: #114) based on the ADMFCA: (a) original CT scan image, (b) ground truth of (a), (c) preprocessed CT scan image of (a), (d) TECI, (e) FSCI, (f) TAECI, (g) FAEI, (h) FCI, (i) histogram of (c), and (j) histogram of (h).
Fig. 7. Clustering of CT scan images (Group: #1, Labels: 1–5) of COVID-19 using the existing clustering algorithms and proposed ADMFCA: (a) preprocessed CT scan image, (b) ground truth, (c) KMC (number of clusters: 3), (d) FCM, (e) GFCM, (f) IIFCM, (g) NEBCA, and (h) ADMFCA.

Fig. 8. Clustering of CT scan images (Group: #1, Labels: 6–10) of COVID-19 using the existing clustering algorithms and proposed ADMFCA: (a) preprocessed CT scan image, (b) ground truth, (c) KMC (number of clusters: 3), (d) FCM, (e) GFCM, (f) IIFCM, (g) NEBCA, and (h) ADMFCA.
Fig. 9. Clustering of CT scan images (Group: #2, Labels: 1–5) of COVID-19 using the existing clustering algorithms and proposed ADMFCA: (a) preprocessed CT scan image, (b) ground truth, (c) KMC (number of clusters: 3), (d) FCM, (e) GFCM, (f) IIFCM, (g) NEBCA, and (h) ADMFCA.

Fig. 10. Clustering of CT scan images (Group: #2, Labels: 6–10) of COVID-19 using the existing clustering algorithms and proposed ADMFCA: (a) preprocessed CT scan image, (b) ground truth, (c) KMC (number of clusters: 3), (d) FCM, (e) GFCM, (f) IIFCM, (g) NEBCA, and (h) ADMFCA.
Fig. 11. Clustering of CT scan images (Group: #3, Labels: 1–5) of COVID-19 using the existing clustering algorithms and proposed ADMFCA: (a) preprocessed CT scan image, (b) ground truth, (c) KMC (number of clusters: 3), (d) FCM, (e) GFCM, (f) IIFCM, (g) NEBCA, and (h) ADMFCA.

Fig. 12. Clustering of CT scan images (Group: #3, Labels: 6–10) of COVID-19 using the existing clustering algorithms and proposed ADMFCA: (a) preprocessed CT scan image, (b) ground truth, (c) KMC (number of clusters: 3), (d) FCM, (e) GFCM, (f) IIFCM, (g) NEBCA, and (h) ADMFCA.
Fig. 13. Clustering of CT scan images (Group: #4, Labels: 1–5) of COVID-19 using the existing clustering algorithms and proposed ADMFCA: (a) preprocessed CT scan image, (b) ground truth, (c) KMC (number of clusters: 3), (d) FCM, (e) GFCM, (f) IIFCM, (g) NEBCA, and (h) ADMFCA.

Fig. 14. Clustering of CT scan images (Group: #4, Labels: 6–10) of COVID-19 using the existing clustering algorithms and proposed ADMFCA: (a) preprocessed CT scan image, (b) ground truth, (c) KMC (number of clusters: 3), (d) FCM, (e) GFCM, (f) IIFCM, (g) NEBCA, and (h) ADMFCA.
Table 4

Performance evaluation of existing clustering algorithms with the proposed ADMFCA with respect to PSNR for clustering the CT scan images of COVID-19.

| Group | Label | CT scan image (preprocessed) | KMC | FCM | GFCM | IIFCM | NEBCA | ADMFCA |
|-------|-------|-------------------------------|-----|-----|------|-------|-------|--------|
| Group #1 | | | | | | | | |
| 1 | #142 | 22.03 | 23.28 | 24.32 | 25.59 | 27.36 | 45.06 |
| 2 | #94 | 21.70 | 23.25 | 24.51 | 25.72 | 27.41 | 43.30 |
| 3 | #105 | 21.46 | 22.18 | 24.36 | 27.16 | 27.52 | 43.30 |
| 4 | #85 | 21.20 | 22.21 | 24.42 | 27.12 | 27.48 | 48.04 |
| 5 | #100 | 22.01 | 22.17 | 24.25 | 26.73 | 27.55 | 45.06 |
| 6 | #110 | 22.01 | 22.17 | 24.25 | 26.73 | 27.55 | 45.06 |
| 7 | #94 | 21.03 | 23.20 | 24.35 | 26.99 | 27.44 | 47.96 |
| 8 | #96 | 21.92 | 23.10 | 24.90 | 26.92 | 27.52 | 47.96 |
| 9 | #109 | 21.73 | 23.19 | 24.75 | 26.99 | 27.53 | 47.92 |
| 10 | #155 | 21.47 | 23.18 | 24.79 | 26.77 | 27.53 | 45.01 |
| Group #2 | | | | | | | | |
| 1 | #118 | 21.47 | 23.18 | 24.79 | 26.77 | 27.53 | 45.01 |
| 2 | #106 | 21.16 | 23.19 | 24.78 | 27.08 | 27.48 | 48.04 |
| 3 | #81 | 21.84 | 23.15 | 24.36 | 27.13 | 27.52 | 43.35 |
| 4 | #71 | 21.90 | 23.13 | 24.34 | 27.12 | 27.51 | 48.00 |
| 5 | #76 | 20.39 | 23.11 | 24.78 | 27.23 | 27.48 | 48.00 |
| 6 | #87 | 20.40 | 23.15 | 24.39 | 27.13 | 27.52 | 48.00 |
| 7 | #113 | 20.40 | 23.15 | 24.39 | 27.13 | 27.52 | 48.00 |
| 8 | #120 | 21.59 | 23.11 | 24.36 | 27.12 | 27.51 | 47.96 |
| 9 | #90 | 21.74 | 23.18 | 24.46 | 26.90 | 27.61 | 47.92 |
| 10 | #179 | 21.20 | 23.15 | 24.39 | 27.10 | 27.45 | 47.96 |
| Group #3 | | | | | | | | |
| 1 | #129 | 21.94 | 23.20 | 24.33 | 26.82 | 27.36 | 47.96 |
| 2 | #97 | 21.79 | 23.19 | 24.51 | 27.29 | 27.71 | 47.92 |
| 3 | #92 | 20.62 | 23.32 | 24.36 | 27.05 | 27.59 | 47.88 |
| 4 | #95 | 20.47 | 23.13 | 24.80 | 27.19 | 27.52 | 47.92 |
| 5 | #87 | 21.05 | 22.82 | 24.86 | 27.02 | 27.55 | 47.92 |
| 6 | #98 | 21.36 | 23.19 | 23.87 | 27.02 | 27.47 | 48.00 |
| 7 | #102 | 21.99 | 23.20 | 24.73 | 27.09 | 27.41 | 45.01 |
| 8 | #107 | 21.95 | 23.02 | 24.32 | 26.95 | 27.67 | 43.35 |
| 9 | #100 | 21.74 | 23.05 | 24.39 | 26.70 | 27.38 | 47.96 |
| 10 | #166 | 21.97 | 23.05 | 24.24 | 26.77 | 27.53 | 45.08 |
| Group #4 | | | | | | | | |
| 1 | #136 | 21.84 | 23.06 | 24.33 | 27.02 | 27.37 | 45.03 |
| 2 | #86 | 21.98 | 22.92 | 24.49 | 27.15 | 27.71 | 45.06 |
| 3 | #100 | 21.46 | 23.18 | 24.18 | 27.05 | 27.59 | 42.06 |
| 4 | #77 | 20.47 | 22.91 | 24.40 | 26.70 | 27.52 | 48.04 |
| 5 | #80 | 21.04 | 22.79 | 24.47 | 27.08 | 27.55 | 48.09 |
| 6 | #92 | 21.34 | 22.92 | 24.21 | 26.88 | 27.47 | 47.88 |
| 7 | #89 | 21.96 | 23.02 | 24.73 | 26.69 | 27.40 | 48.04 |
| 8 | #114 | 21.92 | 23.16 | 24.34 | 26.69 | 27.67 | 45.01 |
| 9 | #104 | 21.73 | 22.79 | 24.39 | 26.96 | 27.38 | 48.00 |
| 10 | #171 | 21.97 | 22.92 | 24.81 | 26.78 | 27.56 | 45.01 |
| Average | | | | | | | | |
| | | 21.49 | 23.03 | 24.47 | 26.89 | 27.51 | 46.59 |
Table 5

Performance evaluation of existing clustering algorithms with the proposed ADMFCA with respect to DSC for clustering the CT scan images of COVID-19.

| Group | Label | CT scan image (preprocessed) | KMC | FCM | GFCM | IIFCM | NEBCA | ADMFCA |
|-------|-------|-------------------------------|-----|-----|------|-------|-------|--------|
| Group #1 |       |                               |     |     |      |       |       |        |
| 1     | #142  | 0.44                          | 0.50| 0.65| 0.78 | 0.82  | 0.90  |        |
| 2     | #94   | 0.40                          | 0.51| 0.66| 0.76 | 0.81  | 0.91  |        |
| 3     | #105  | 0.37                          | 0.54| 0.68| 0.77 | 0.83  | 0.89  |        |
| 4     | #85   | 0.39                          | 0.52| 0.65| 0.78 | 0.84  | 0.93  |        |
| 5     | #100  | 0.40                          | 0.53| 0.63| 0.77 | 0.83  | 0.94  |        |
| 6     | #110  | 0.44                          | 0.49| 0.67| 0.75 | 0.84  | 0.92  |        |
| 7     | #94   | 0.45                          | 0.58| 0.64| 0.78 | 0.85  | 0.93  |        |
| 8     | #96   | 0.39                          | 0.53| 0.63| 0.77 | 0.83  | 0.90  |        |
| 9     | #109  | 0.39                          | 0.54| 0.64| 0.78 | 0.84  | 0.94  |        |
| 10    | #155  | 0.39                          | 0.59| 0.63| 0.78 | 0.84  | 0.87  |        |
| Group #2 |       |                               |     |     |      |       |       |        |
| 1     | #118  | 0.39                          | 0.49| 0.68| 0.78 | 0.90  | 0.91  |        |
| 2     | #106  | 0.39                          | 0.49| 0.65| 0.77 | 0.85  | 0.93  |        |
| 3     | #81   | 0.40                          | 0.51| 0.64| 0.78 | 0.83  | 0.91  |        |
| 4     | #71   | 0.45                          | 0.50| 0.62| 0.75 | 0.87  | 0.93  |        |
| 5     | #76   | 0.39                          | 0.49| 0.65| 0.78 | 0.83  | 0.94  |        |
| 6     | #87   | 0.40                          | 0.51| 0.66| 0.75 | 0.89  | 0.91  |        |
| 7     | #113  | 0.39                          | 0.52| 0.64| 0.77 | 0.86  | 0.93  |        |
| 8     | #120  | 0.42                          | 0.58| 0.65| 0.76 | 0.81  | 0.93  |        |
| 9     | #90   | 0.45                          | 0.47| 0.66| 0.78 | 0.84  | 0.92  |        |
| 10    | #179  | 0.43                          | 0.46| 0.64| 0.77 | 0.85  | 0.92  |        |
| Group #3 |       |                               |     |     |      |       |       |        |
| 1     | #129  | 0.37                          | 0.49| 0.66| 0.79 | 0.84  | 0.91  |        |
| 2     | #97   | 0.39                          | 0.52| 0.65| 0.78 | 0.85  | 0.93  |        |
| 3     | #92   | 0.42                          | 0.54| 0.66| 0.74 | 0.83  | 0.89  |        |
| 4     | #95   | 0.45                          | 0.54| 0.64| 0.76 | 0.86  | 0.93  |        |
| 5     | #87   | 0.39                          | 0.52| 0.62| 0.75 | 0.84  | 0.92  |        |
| 6     | #98   | 0.39                          | 0.59| 0.66| 0.77 | 0.86  | 0.89  |        |
| 7     | #102  | 0.38                          | 0.48| 0.65| 0.75 | 0.84  | 0.94  |        |
| 8     | #107  | 0.37                          | 0.53| 0.66| 0.77 | 0.82  | 0.89  |        |
| 9     | #100  | 0.37                          | 0.47| 0.66| 0.74 | 0.84  | 0.93  |        |
| 10    | #166  | 0.37                          | 0.49| 0.64| 0.77 | 0.85  | 0.91  |        |
| Group #4 |       |                               |     |     |      |       |       |        |
| 1     | #136  | 0.37                          | 0.50| 0.63| 0.79 | 0.87  | 0.92  |        |
| 2     | #86   | 0.38                          | 0.48| 0.65| 0.78 | 0.84  | 0.92  |        |
| 3     | #100  | 0.45                          | 0.56| 0.64| 0.79 | 0.82  | 0.86  |        |
| 4     | #77   | 0.39                          | 0.58| 0.63| 0.76 | 0.88  | 0.93  |        |
| 5     | #80   | 0.40                          | 0.47| 0.66| 0.78 | 0.94  | 0.92  |        |
| 6     | #92   | 0.46                          | 0.57| 0.65| 0.77 | 0.87  | 0.92  |        |
| 7     | #89   | 0.46                          | 0.49| 0.64| 0.78 | 0.88  | 0.94  |        |
| 8     | #114  | 0.45                          | 0.57| 0.63| 0.79 | 0.86  | 0.87  |        |
| 9     | #104  | 0.46                          | 0.48| 0.64| 0.77 | 0.83  | 0.94  |        |
| 10    | #171  | 0.38                          | 0.47| 0.65| 0.78 | 0.86  | 0.92  |        |
| Average|       |                               | 0.41| 0.52| 0.65 | 0.77  | 0.85  | 0.92  |

Fig. 15. Comparison of MSE values with the existing clustering algorithms and the proposed ADMFCA.

yields an average DSC value of 0.92, which is significantly higher than the existing clustering methods. This high DSC value indicates that the FCIs produced by the proposed ADMFCA are identical to their respective ground truths.

- The JSC values for the existing competing methods and the proposed ADMFCA are shown in Table 6. The proposed ADMFCA obtains an average JSC value of 0.96 for the FCIs, which is significantly higher than the existing clustering methods. For the proposed ADMFCA, the high JSC value
Table 6
Performance evaluation of existing clustering algorithms with the proposed ADMFCA with respect to JSC for clustering the CT scan images of COVID-19.

| Group   | Label | CT scan image (preprocessed) | KMC | FCM | GFCM | IIFCM | NEBCA | ADMFCA |
|---------|-------|------------------------------|-----|-----|------|-------|-------|--------|
| Group #1|       |                              |     |     |      |       |       |        |
| 1       | #142  | 0.48                         | 0.54| 0.69| 0.82 | 0.86  | 0.94  |        |
| 2       | #94   | 0.44                         | 0.55| 0.70| 0.80 | 0.85  | 0.95  |        |
| 3       | #105  | 0.41                         | 0.58| 0.72| 0.81 | 0.87  | 0.93  |        |
| 4       | #85   | 0.43                         | 0.56| 0.69| 0.82 | 0.88  | 0.97  |        |
| 5       | #100  | 0.44                         | 0.57| 0.67| 0.81 | 0.87  | 0.98  |        |
| 6       | #110  | 0.48                         | 0.53| 0.71| 0.79 | 0.83  | 0.96  |        |
| 7       | #94   | 0.49                         | 0.62| 0.68| 0.82 | 0.89  | 0.97  |        |
| 8       | #96   | 0.43                         | 0.57| 0.67| 0.81 | 0.87  | 0.94  |        |
| 9       | #109  | 0.43                         | 0.58| 0.68| 0.82 | 0.88  | 0.98  |        |
| 10      | #155  | 0.43                         | 0.63| 0.67| 0.82 | 0.88  | 0.91  |        |
| Group #2|       |                              |     |     |      |       |       |        |
| 1       | #118  | 0.43                         | 0.53| 0.72| 0.82 | 0.94  | 0.95  |        |
| 2       | #106  | 0.43                         | 0.53| 0.69| 0.81 | 0.89  | 0.97  |        |
| 3       | #81   | 0.40                         | 0.51| 0.64| 0.78 | 0.83  | 0.91  |        |
| 4       | #71   | 0.49                         | 0.54| 0.66| 0.79 | 0.91  | 0.97  |        |
| 5       | #76   | 0.43                         | 0.53| 0.69| 0.82 | 0.87  | 0.98  |        |
| 6       | #87   | 0.44                         | 0.55| 0.70| 0.79 | 0.93  | 0.95  |        |
| 7       | #113  | 0.43                         | 0.56| 0.68| 0.81 | 0.90  | 0.97  |        |
| 8       | #120  | 0.46                         | 0.62| 0.69| 0.80 | 0.85  | 0.97  |        |
| 9       | #90   | 0.49                         | 0.51| 0.70| 0.82 | 0.88  | 0.96  |        |
| 10      | #179  | 0.47                         | 0.50| 0.68| 0.81 | 0.89  | 0.96  |        |
| Group #3|       |                              |     |     |      |       |       |        |
| 1       | #129  | 0.41                         | 0.53| 0.70| 0.83 | 0.88  | 0.95  |        |
| 2       | #97   | 0.43                         | 0.56| 0.69| 0.82 | 0.89  | 0.97  |        |
| 3       | #92   | 0.46                         | 0.58| 0.70| 0.78 | 0.87  | 0.93  |        |
| 4       | #95   | 0.49                         | 0.58| 0.68| 0.80 | 0.90  | 0.97  |        |
| 5       | #87   | 0.43                         | 0.56| 0.66| 0.79 | 0.88  | 0.96  |        |
| 6       | #98   | 0.43                         | 0.63| 0.70| 0.81 | 0.90  | 0.93  |        |
| 7       | #102  | 0.42                         | 0.52| 0.69| 0.79 | 0.88  | 0.98  |        |
| 8       | #107  | 0.41                         | 0.57| 0.70| 0.81 | 0.86  | 0.93  |        |
| 9       | #100  | 0.41                         | 0.51| 0.70| 0.78 | 0.88  | 0.97  |        |
| 10      | #166  | 0.41                         | 0.53| 0.68| 0.81 | 0.89  | 0.95  |        |
| Group #4|       |                              |     |     |      |       |       |        |
| 1       | #136  | 0.41                         | 0.54| 0.67| 0.83 | 0.91  | 0.96  |        |
| 2       | #86   | 0.42                         | 0.52| 0.69| 0.82 | 0.88  | 0.96  |        |
| 3       | #100  | 0.49                         | 0.60| 0.68| 0.83 | 0.86  | 0.90  |        |
| 4       | #77   | 0.43                         | 0.62| 0.67| 0.80 | 0.92  | 0.97  |        |
| 5       | #80   | 0.44                         | 0.51| 0.70| 0.82 | 0.90  | 0.96  |        |
| 6       | #92   | 0.50                         | 0.61| 0.69| 0.81 | 0.91  | 0.96  |        |
| 7       | #89   | 0.50                         | 0.53| 0.68| 0.82 | 0.92  | 0.98  |        |
| 8       | #114  | 0.50                         | 0.53| 0.68| 0.82 | 0.92  | 0.98  |        |
| 9       | #104  | 0.50                         | 0.52| 0.68| 0.81 | 0.87  | 0.98  |        |
| 10      | #171  | 0.42                         | 0.51| 0.69| 0.82 | 0.90  | 0.96  |        |
| Average | –     | 0.45                         | 0.56| 0.69| 0.81 | 0.89  | 0.96  |        |

Fig. 16. Comparison of PSNR values with the existing clustering algorithms and the proposed ADMFCA.

indicates that the regions of interest of the FCIs are almost identical to their respective ground truths.

- Table 7 summarizes the CC values obtained from existing methods and the proposed ADMFCA. The average CC values for KMC, FCM, GFCM, IIFCM, NEBCA and the proposed ADMFCA are 0.46, 0.57, 0.70, 0.82, 0.90 and 0.97, respectively. This CC value of the proposed ADMFCA suggests that the clustered grayscale values of the FCIs are strongly similar to their respective ground truths.

Comparison curves for various MSE, PSNR, DSC, JSC and CC (listed in Tables 3–7) are plotted and shown in Figs. 15–19, respectively. These comparison curves show that the proposed ADMFCA outperforms the existing clustering methods (i.e., KMC, FCM, GFCM, IIFCM and NEBCA) in terms of MSE, PSNR, DSC,
Table 7
Performance evaluation of existing clustering algorithms with the proposed ADMFCA with respect to CC for clustering the CT scan images of COVID-19.

| Group | Label | CT scan image (preprocessed) | KMC | FCM | GFCM | IIFCM | NEBCA | ADMFCA |
|-------|-------|-----------------------------|-----|-----|------|-------|-------|--------|
|       |       |                             |     |     |      |       |       |        |
|       | 1     | #142                        | 0.49| 0.55| 0.70 | 0.83  | 0.87  | 0.95   |
|       | 2     | #94                         | 0.45| 0.56| 0.71 | 0.81  | 0.86  | 0.96   |
|       | 3     | #105                        | 0.42| 0.59| 0.73 | 0.82  | 0.88  | 0.94   |
|       | 4     | #85                         | 0.44| 0.57| 0.70 | 0.83  | 0.89  | 0.98   |
|       | 5     | #100                        | 0.45| 0.58| 0.68 | 0.82  | 0.88  | 0.99   |
| Group #1 |    |                              |     |     |      |       |       |        |
|       | 1     | #118                        | 0.44| 0.54| 0.73 | 0.83  | 0.95  | 0.96   |
|       | 2     | #106                        | 0.44| 0.54| 0.70 | 0.82  | 0.90  | 0.98   |
|       | 3     | #81                         | 0.45| 0.56| 0.69 | 0.83  | 0.88  | 0.96   |
|       | 4     | #71                         | 0.50| 0.55| 0.67 | 0.80  | 0.92  | 0.98   |
|       | 5     | #76                         | 0.44| 0.54| 0.70 | 0.83  | 0.88  | 0.99   |
|       | 6     | #87                         | 0.45| 0.56| 0.71 | 0.80  | 0.94  | 0.96   |
|       | 7     | #113                        | 0.44| 0.57| 0.69 | 0.82  | 0.91  | 0.98   |
|       | 8     | #120                        | 0.47| 0.63| 0.70 | 0.81  | 0.86  | 0.98   |
|       | 9     | #90                         | 0.50| 0.52| 0.71 | 0.83  | 0.89  | 0.97   |
|       | 10    | #179                        | 0.48| 0.51| 0.69 | 0.82  | 0.90  | 0.97   |
| Group #2 |    |                              |     |     |      |       |       |        |
|       | 1     | #129                        | 0.42| 0.54| 0.71 | 0.84  | 0.89  | 0.96   |
|       | 2     | #97                         | 0.44| 0.57| 0.70 | 0.83  | 0.90  | 0.98   |
|       | 3     | #92                         | 0.47| 0.59| 0.71 | 0.79  | 0.88  | 0.94   |
|       | 4     | #95                         | 0.50| 0.59| 0.69 | 0.81  | 0.91  | 0.98   |
|       | 5     | #87                         | 0.44| 0.57| 0.67 | 0.80  | 0.89  | 0.97   |
|       | 6     | #98                         | 0.44| 0.64| 0.71 | 0.82  | 0.91  | 0.94   |
|       | 7     | #102                        | 0.43| 0.53| 0.70 | 0.80  | 0.89  | 0.99   |
|       | 8     | #107                        | 0.42| 0.58| 0.71 | 0.82  | 0.87  | 0.94   |
|       | 9     | #100                        | 0.42| 0.52| 0.71 | 0.79  | 0.89  | 0.98   |
|       | 10    | #166                        | 0.42| 0.54| 0.69 | 0.82  | 0.90  | 0.96   |
| Group #3 |    |                              |     |     |      |       |       |        |
|       | 1     | #136                        | 0.42| 0.55| 0.68 | 0.84  | 0.92  | 0.97   |
|       | 2     | #86                         | 0.43| 0.53| 0.70 | 0.83  | 0.89  | 0.97   |
|       | 3     | #100                        | 0.50| 0.61| 0.69 | 0.84  | 0.87  | 0.91   |
|       | 4     | #77                         | 0.44| 0.63| 0.68 | 0.81  | 0.93  | 0.98   |
|       | 5     | #80                         | 0.45| 0.52| 0.71 | 0.83  | 0.89  | 0.97   |
|       | 6     | #92                         | 0.51| 0.62| 0.70 | 0.82  | 0.92  | 0.97   |
|       | 7     | #89                         | 0.51| 0.54| 0.69 | 0.83  | 0.93  | 0.99   |
|       | 8     | #114                        | 0.50| 0.62| 0.68 | 0.84  | 0.91  | 0.92   |
|       | 9     | #104                        | 0.51| 0.53| 0.69 | 0.82  | 0.88  | 0.99   |
|       | 10    | #171                        | 0.43| 0.52| 0.70 | 0.83  | 0.91  | 0.97   |
| Average |     |                             | 0.46| 0.57| 0.70 | 0.82  | 0.90  | 0.97   |

Fig. 17. Comparison of DSC values with the existing clustering algorithms and the proposed ADMFCA.

JSC and CC for clustering the CT scan images of COVID-19 patients. Consequently, the proposed ADMFCA is highly effective at forming clusters of pixels associated with infected regions.

7. Conclusions and future directions

In this study, ambiguous set theory was discussed, which was recently proposed to address inherent uncertainties of events. The ambiguous set theory can be considered as an extension of three existing theories, viz., fuzzy set, intuitionistic fuzzy set and neutrosophic set. The main robustness of this theory was its ability to represent the ambiguity of an uncertain event with four distinct degree of memberships, called true, false, true-ambiguous and false-ambiguous. To endorse this theory, various definitions,
Formulas and properties were discussed in this study. The main contributions of this study are summarized as:

- To measure the ambiguity associated with four degree of memberships, four different entropies were defined, called TE, FE, TAE and FAE.
- This study proposed a new image clustering algorithm using the concepts of ambiguous set, entropies (TE, FE, TAE and FAE) and image fusion, called ADMFCA.
- The primary application of the proposed ADMFCA was illustrated in the clustering of chest CT images of COVID-19 patients. This algorithm was allowed to generate four different clustered images based on specified number of clusters. These four clustered images were referred to as TECI, FECI, TAECl and FAECI. Finally, FCI's were generated by combining TECI, FECI, TAECl and FAECI using the image fusion operation. The main purpose of image fusion was to include the best features of TECI, FECI, TAECl and FAECI into FCIs.
- The performance of the proposed ADMFCA was compared against existing clustering methods, including KMC, FCM, FCM, FCM and NEBCA. Various performance evaluation metrics, such as MSE, PSNR, DSC, JSC and CC indicated that the proposed ADFMCA outperformed the existing clustering methods.

It can be concluded that the proposed ADMFCA was proven to be effective in clustering CT scan images of COVID-19 patients. Therefore, the proposed ADMFCA can be considered as a new promising diagnostic method for health professionals. The main limitation of the study was that the proposed ADMFCA was validated only on chest CT scan images of COVID-19 patients. In the future, the proposed ADMFCA can be verified and validated with other forms of digital images, such as X-rays, MRIs [40], remotely sensed high-resolution satellite images [48], and so on. Additionally, the proposed ADMFCA can be used to cluster a variety of numerical data, including meteorological data, financial data, stock market data, and so on.

**Ethical approval**

This article does not contain any studies with human participants performed by any of the authors.

**CRediT authorship contribution statement**

Pritpal Singh: Conceptualization, Software, Methodology, Theories and definitions, Mathematical deductions and their proofs, Writing - review & editing. Surya Sekhar Bose: Conceptualization, Software, Methodology, Theories and definitions, Mathematical deductions and their proofs, Writing - review & editing.

**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.

**References**

[1] Y.Y. Zheng, et al., Covid-19 and the cardiovascular system, Nat. Rev. Cardiol. 17 (2020) 259–260.

[2] P. Singh, S.S. Bose, A quantum-clustering optimization method for COVID-19 CT scan image segmentation, Expert Syst. Appl. 185 (2021) 115637.
[26] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.

[27] K.T. Atanassov, Intuitionistic fuzzy sets, in: Proc. VII IFIP's Session, Sofia, Bulgaria, 1983, pp. 16840–1697.

[28] F. Smarandache, Neutrosophy, a new branch of philosophy, Infinite Study (2002).

[29] L. Ma, R. Staunton, A modified fuzzy c-means image segmentation algorithm for use with uneven illumination patterns, Pattern Recognit. 40 (2007) 3005–3011.

[30] P. Qian, et al., Knowledge-leveraged transfer fuzzy C-Means for image segmentation, Knowl.-Based Syst. 130 (2017) 33–50.

[31] L. Chen, C.L.P. Chen, M. Lu, A multiple-kernel fuzzy c-means algorithm for image segmentation, IEEE Trans. Syst. Man Cybern. 41 (2011) 1263–1274.

[32] Z. Ma, Y. Xiao, Q. Chen, Q. Sun, D. Xia, D.D. Feng, Fuzzy c-means clustering with weighted image patch for image segmentation, Appl. Soft Comput. 12 (2012) 1659–1667.

[33] Z. Wang, Q. Song, Y.C. Soh, K. Sim, An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Underst. 117 (2013) 1412–1420.

[34] F. Zhao, J. Liao, H. Liu, Kernel generalized fuzzy c-means clustering with spatial information for image segmentation, Digit. Signal Process. 23 (2013) 184–199.

[35] T. Chaiera, A novel intuitionistic fuzzy c means clustering algorithm and its application to medical images, Appl. Soft Comput. 11 (2011) 1711–1717.

[36] J. Han, S. Yang, B. Lee, A novel 3-D color histogram equalization method, IEEE Trans. Circuits Syst. Video Technol. 11 (2001) 379–386.

[37] S. Varela-Santos, P. Melin, A novel medical diagnosis model for COVID-19 infection detection based on deep features and bayesian optimization, Appl. Soft Comput. 97 (2020) 106580.

[38] H. Kang, et al., Diagnosis of coronavirus disease 2019 (COVID-19) with structured latent multi-view representation learning, IEEE Trans. Med. Imag. 39 (2020) 2615–2625.

[39] S. Wang, et al., A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT, IEEE Trans. Med. Imag. 39 (2020) 2606–2614.

[40] M. Nour, Z. Comert, K. Polat, A novel medical diagnosis model for COVID-19 infection detection based on deep features and bayesian optimization, Appl. Soft Comput. 97 (2020) 106580.

[41] H. Kang, et al., Diagnosis of coronavirus disease 2019 (COVID-19) with structured latent multi-view representation learning, IEEE Trans. Med. Imag. 39 (2020) 2615–2625.

[42] X. Wang, et al., A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT, IEEE Trans. Med. Imag. 39 (2020) 2615–2625.

[43] S. Varela-Santos, P. Melin, A new approach for classifying coronavirus COVID-19 based on its manifestation on chest X-rays using texture features and neural networks, Inform. Sci. 545 (2021) 403–414.

[44] M. Tang, H. Liao, Managing information measures for hesitant fuzzy linguistic term sets and their applications in designing clustering algorithms, Inf. Fusion 50 (2019) 30–42.

[45] P. Qian, et al., Knowledge-leveraged transfer fuzzy C-Means for texture image segmentation, Knowl.-Based Syst. 130 (2017) 33–50.

[46] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.

[47] K.T. Atanassov, Intuitionistic fuzzy sets, in: Proc. VII IFIP's Session, Sofia, Bulgaria, 1983, pp. 16840–1697.

[48] F. Smarandache, Neutrosophy, a new branch of philosophy, Infinite Study (2002).

[49] L. Ma, R. Staunton, A modified fuzzy c-means image segmentation algorithm for use with uneven illumination patterns, Pattern Recognit. 40 (2007) 3005–3011.

[50] P. Qian, et al., Knowledge-leveraged transfer fuzzy C-Means for image segmentation, Knowl.-Based Syst. 130 (2017) 33–50.

[51] L. Chen, C.L.P. Chen, M. Lu, A multiple-kernel fuzzy c-means algorithm for image segmentation, IEEE Trans. Syst. Man Cybern. 41 (2011) 1263–1274.

[52] Z. Ji, Y. Xiao, Q. Chen, Q. Sun, D. Xiang, D.D. Feng, Fuzzy c-means clustering with weighted image patch for image segmentation, Appl. Soft Comput. 12 (2012) 1659–1667.

[53] Z. Wang, Q. Song, Y.C. Soh, K. Sim, An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Underst. 117 (2013) 1412–1420.

[54] F. Zhao, J. Liao, H. Liu, Kernel generalized fuzzy c-means clustering with spatial information for image segmentation, Digit. Signal Process. 23 (2013) 184–199.

[55] T. Chaiera, A novel intuitionistic fuzzy c-means clustering algorithm and its application to medical images, Appl. Soft Comput. 11 (2011) 1711–1717.

[56] J. Han, S. Yang, B. Lee, A novel 3-D color histogram equalization method, IEEE Trans. Circuits Syst. Video Technol. 11 (2001) 379–386.

[57] S. Varela-Santos, P. Melin, A novel medical diagnosis model for COVID-19 infection detection based on deep features and bayesian optimization, Appl. Soft Comput. 97 (2020) 106580.

[58] H. Kang, et al., Diagnosis of coronavirus disease 2019 (COVID-19) with structured latent multi-view representation learning, IEEE Trans. Med. Imag. 39 (2020) 2615–2625.

[59] X. Wang, et al., A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT, IEEE Trans. Med. Imag. 39 (2020) 2615–2625.

[60] M. Nour, Z. Comert, K. Polat, A novel medical diagnosis model for COVID-19 infection detection based on deep features and bayesian optimization, Appl. Soft Comput. 97 (2020) 106580.