Mansoura Engineering Journal

Volume 49 | Issue 3 Article 18

2024

An Innovative COVID-19 Patient Recognition Framework

Shereen H. Ali
Communications & Electronics Engineering Department, Delta Higher Institute for Engineering & Technology, Mansoura, Egypt, engshereen2011@gmail.com

ORIGINAL STUDY

An Innovative COVID-19 Patient Recognition Framework

Shereen H. Ali

Department of Communications and Electronics Engineering, Delta Higher Institute for Engineering and Technology, Mansoura, Egypt

Abstract

A virus called coronavirus disease 2019 (COVID-19) has caused devastation throughout the world and is still putting people's lives in jeopardy. It is important to identify COVID-19 patients as soon as possible so that they can be treated and kept from spreading. A new framework for recognizing COVID-19-infected individuals would be provided in this study. The patient's recognition framework (PRF) is a term used to describe a method for detecting patients. The PRF consists of three stages, which are: the pre-processing stage (P^2S), the feature selection stage, and the classification stage. The P^2S extracts a collection of features from computed tomography chest scan images for a variety of people, some of whom are infected with COVID-19 and others who are not. Feature selection stage selects only the most beneficial characteristics when detecting COVID-19 patients in P^2S by using the enhanced moth flame optimization approach as a wrapper method. The classification stage employs the support vector machine classifier to accurately detect COVID-19-contaminated individuals with the shortest possible time cost, relying on enhanced moth flame optimization's significant features. According to analytical outcomes, the PRF strategy exceeds contemporary techniques in terms of efficiency. PRF achieves the highest accuracy, precision, and recall. Besides, it achieves the lowest error, with values equal to 98, 93, 92, and 2%, respectively.

Keywords: Coronavirus disease 2019, Feature selections, Moth flame optimization, Rough set

1. Introduction

The novel coronavirus, also known as COVID-19, is rapidly propagating around the world, posing a threat to public health (Sheela and Arun, 2022; Chung et al., 2019). As of March 2020, COVID-19 is considered an international pandemic by the World Health Organization (WHO) https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020. COVID-19 is essentially a contagious virus that spreads through droplets; either explicit transmission occurs when a patient or disease carrier coughs or sneezes, or implicit transmission occurs when a person shakes hands, uses personal items, or touches surfaces that have been smeared with the virus's droplets. The mouth, nose, and eye mucous membranes are the entry points for the virus into the human body (Ghose et al., 2022; Mahanty et al., 2022; Zhu et al., 2022). The symptoms of this illness include fever, headache, muscle soreness, runny nose, cough, sore throat, and lung infection (Huang et al., 2020). It is detrimental to everyday living, public health, and the global economy (Kuzmenko et al., 2023). Furthermore, COVID-19 infections demolish the healthcare system in less than four weeks after they begin to spread (Zaim et al., 2020). Quick and precise COVID-19 detection is becoming increasingly important for limiting contamination and helping patients avoid disease processes. Real-time polymerase chain reaction (RT-PCR) testing in laboratories is the most widely used diagnostic technique available today. However, it takes a lot of time and money (in some countries) (Iancu et al., 2022). Radiography and computed tomography (CT) scans have a major role in the early detection and identification of COVID-19 cases (Akl et al., 2023).

Data mining is a useful technique that may be applied to predict medical conditions and provide

Received 30 October 2023; revised 6 February 2024; accepted 24 February 2024.
Available online 22 April 2024
E-mail address: Engshereen2011@gmail.com.

https://doi.org/10.58491/2735-4202.3201

healthcare providers with the necessary knowledge to make well-informed decisions (Vommi and Battula, 2023; Swathy and Saruladha, 2022; Alenezi and Alqenaei, 2021). Since it completes difficult computing tasks like identifying patterns in large datasets, it can be used to identify and obtain relevant patterns for diagnostic categorization were categorized using classification, which is one of the analytical procedures.

Support vector machines (SVMs), one of the most popular classifiers, may be extended to other fields and tolerate high-dimensional data since they are based on the notion of a theory of statistical learning (Vapnik, 1999; Narin, 2020). Many complex application domains, including clinical evaluation (Narin, 2020), network intrusion detection (Ruoyuan et al., 2022), and weather prediction (Chengcheng et al., 2022), have demonstrated the efficacy of SVM. The rapid development and severity of COVID-19 make it imperative to create accurate and timely diagnostic methods in order to combat the disease as soon as possible. In fact, COVID-19 patients can be identified using classification algorithms based on information extracted from CT images. CT scans must undergo feature extraction before a prediction method is applied. Retrieving the features from a CT scan is the main objective of feature extraction, as this allows the classification system to use the features to draw accurate judgements (Chiranji and Acharjya).

Wavelet transforms, Gabor filters, co-occurrence matrixes, and other methods can all be used to extract features (Mohan and Subashini, 2018; Kanagaraj and Kumar, 2020). In actuality, the Gray Level Co-occurrence Matrix (GLCM) is the most effective method for characterizing the texture of an image in image analysis methods (Zotin et al., 2019). There could be a variety of unnecessary or redundant characteristics in the retrieved features. Therefore, removing those unnecessary features is a procedure that is required before beginning to learn the classification algorithm. By choosing the important features, the classification approach may correctly categorize COVID-19 individuals with the least amount of time lost. There are many different feature selection methods categorized as filter, wrapper, and hybrid approaches (Sadeghian et al., 2021; Wah et al., 2018).

Medicine is one of the most productive and emerging fields in feature selection and machine learning applications, with the goal of reducing not only the dimensionality of problems but also the costs involved; for example, extracting information from images or understanding the reasons for disagreements among image-analysis experts regarding disease diagnosis (Naheed et al., 2020). One of the strategies for dimensionality reduction is feature selection, in which relevant features are chosen while irrelevant and redundant features are removed (Sadeghian et al., 2021). Reduced input dimensionality can boost performance by lowering learning speed and model complexity while also improving generalization capacity and classification accuracy. The right features can also lead to significant savings on measurements and a better understanding of the situation.

Moth Flame Optimization (MFO) (Mirjalili, 2015) is a meta-heuristic technique that stimulates and mimics the natural behavior of moths. Moths fly in a straight line by defining an angle around the moon and converge on its light. Meta-heuristic optimizers tend to explore a range of solutions in a way that heuristically differs from one algorithm to another and mimics a natural observation, starting by exploring the search space stage where the near-best solutions are highlighted and starting to focus search in such areas, which is called the 'exploitation phase.' As an optimization method, MFO offers many benefits, including: (i) having few setting parameters; (ii) being easy to understand and implement; and (iii) having fast convergence. As a result, MFO can be used successfully as a wrapper feature selection approach.

The main objective of this research is to offer an accurate COVID-19 infection recognition framework that takes the least amount of time from a chest CT image by integrating the strengths of various techniques. The proposed framework is divided into three stages, namely, the pre-processing stage (P^2S), the feature selection stage (FS^2), and the classification stage (CS). Using the GLCM approach, a combination of features was retrieved from CT scans during P^2S. The proposed enhanced moth flame optimization (EMFO) approach is utilized in FS^2 to determine the most significant features from those retrieved in P^2S for the subsequent CS. The SVM classifier is utilized in the third stage CS to provide quick and exact recognition of COVID-19-infected patients using specified features from FS^2. The experimental results show that the proposed framework is capable of accurately estimating the optimal features while maintaining a quick convergence rate. Furthermore, the identification of COVID-19 patients has high accuracy. Some of the main contributions in this study are described as follows:

(1) A new COVID-19 patient recognition framework (PRF) for precise recognition of infected individuals was thus developed.
(2) The PRF is divided into three stages: (i) the P^2S, (ii) the FS^2 and (iii) the CS.

(3) By selecting a feature subset from CT scans using the GLCM technique, the datasets on COVID-19 patients and non-COVID-19 individuals are gathered in the P^2S and then displayed in a way that is appropriate for the recognition model.
(4) During FS^2, the most significant features will be picked from P^2S by utilizing the suggested EMFO approach to provide useful data patterns during CS.
(5) During CS, the SVM classifier enables rapid and precise recognition of COVID-19 patients based on specific FS^2 features.
(6) The efficiency of the suggested PRF is examined with the comparison of state-of-the-art methods using a variety of performance metrics such as precision, recall, f1-measure, accuracy, and run time.

The paper is outlined as follows: Section two provides PRF's ability to recognize COVID-19. Section three discusses the previous efforts regarding COVID-19 patient detection. Section four discusses an overview of moth flame optimization, rough set basics, and SVM classifiers. Section five focuses on the proposed PRF. The experiments are presented, and the results are analyzed in Section six. The discussion is presented in Section seven. Finally, the conclusion is summarized in Section eight.

2. PRF's ability to recognize COVID-19

The terrifying COVID-19 viral outbreak resembled a pandemic in its rapid spread. Without a doubt, it represents the worst problem humanity has encountered since Second World War. On the other hand, COVID-19 is far more than just a health issue; it has the opportunity to have devastating social - financial effects that will surely leave an indelible mark (Chung et al., 2019). The three approaches for recognizing COVID-19 are as follows: (1) RT-PCR, (2) chest CT imaging scan, or (3) numerical laboratory testing (NLT). The RT-PCR is now the 'guideline' for determining whether an individual has COVID-19 positivity.

RT-PCR testing is rapid, sensitive, and consistent. Chemicals are used to eliminate any lipids, proteins, or other compounds from a sample taken from a patient's mouth or nose, leaving only RNA (Iancu et al., 2022). The genetic material of an individual and, if any, coronavirus RNA are combined to create isolated RNA. The RT-PCR test is susceptible to mistakes and provides false findings; therefore, it cannot detect all illnesses (Iancu et al., 2022). Consequently, a negative RT-PCR result may not completely rule out COVID-19 disease. Such undetected cases can have devastating consequences because to COVID-19's exponential proliferation. As a result, RT-PCR should not be employed as the sole means of discovering COVID-19 patients (Iancu et al., 2022). Because COVID-19 is easily recognized by cloudy, erratic, 'ground glass' white spots in the lungs, a chest CT scan can diagnose the condition. Several investigations have found that CT has much better sensitivity than RT-PCR in diagnosing COVID-19 (Vommi and Battula, 2023).

The suggested PRF in this paper focuses on data mining methods, more specifically classification, to recognize COVID-19. PRF relies on SVM, a supervised learning classification approach, despite the fact that a variety of classification algorithms can be employed (Vapnik, 1999). Because of the following factors, we assert that SVM is the best appropriate classifier that can be employed for COVID-19 diagnosis as a forecasting model: (i) SVM is easy to use, adaptable, quick, and suitable for scenarios in reality. (ii) SVM can create good predictions even with little training data because it only needs a minimal quantity to estimate the parameters needed to build the classification model. (iii) Since it is appropriate for progressive training, SVM can instantly learn new data. (iv) In addition to retaining the benefits of conventional SVM, PRF has been improved by new feature selection methods. SVM uses a subset of training points in the classification model, making it memory-effective. The adoption of the suggested PRF as the COVID-19 recognition technique, which completely relies on CT chest imaging rather than NLTs or RT-PCR tests, is demonstrated by the deployment of a new feature selection method, as will be seen from the experimental findings.

3. Related work

In Maghdid et al. (2020), based on a dataset of chest radiography and CT images, a convolutional neural network (CNN) model was proposed to identify COVID-19 instances. The two primary mechanisms used in the proposed CNN model are the AlexNet transfer-learning method and the CNN structure. Although this proposed approach is straightforward, its accuracy is insufficient for recognizing COVID-19 patients. The testing findings made it clear that employing a pretrained network produced the models with the highest level of accuracy. Using the modified CNN, the accuracy was increased to 94.1% with a minimal time penalty.

In Sun et al. (2020), to categorize COVID-19 individuals, a deep forest model has been used to learn a high-level representation of features based on chest CT scans. To reduce feature duplication, an adaptive feature selection process was used based

on the trained forest. The experimental findings showed that the proposed model achieved values equal to 91.79, 93.05, 89.95, and 96.35%, respectively, for accuracy, sensitivity, specificity, and area under curve (AUC).

In Nazish et al. (2021), by analyzing chest radiography, authors employed SVM and logistic regression to distinguish COVID-19 from noninfected patients. When compared with logistic regression, the SVM has been observed to have an accuracy of 96%.

In Sen et al. (2021), to discover COVID-19 from chest CT scans, a CNN has been used to select features from chest CT images in the first phase. The feature qualities are ranked using two separate filter methods: mutual information and Relief-F. In the second phase, a feature selection method was used to find the best features for identifying COVID and non-COVID patients from chest CT images. For detection, the resulting feature collection is passed to an SVM classifier. 90% accuracy was reached with the proposed model.

In Zhang et al. (2021), a five-layer deep convolutional neural network with stochastic pooling for chest CT-based COVID-19 diagnosis was introduced. Three improvements were presented: (i) stochastic pooling to replace average pooling and max pooling; (ii) the convolution layer was combined with the batch normalization layer and obtained the convolution block (CB); (iii) the dropout layer was merged with the fully connected layer and obtained the fully connected block (FCB). Experimental results showed that the proposed model is effective in detecting COVID-19 based on chest CT images. A 5L-DCNN-SP-C algorithm achieved a sensitivity of 93.28% ± 1.50%, a specificity of 94.00% ± 1.56%, and an accuracy of 93.64% ± 1.42%. Table 1 provides a succinct summary of recent COVID-19 detection research.

In Habib et al. (2022), the capabilities of machine learning approaches in determining COVID-19 patients were utilized, integrating both traditional methods and investigating the influence of viral infection using chest radiography images. This method consists of preparation, obtaining features, and classification. The histogram of orientation and local binary pattern feature descriptors are used for gathering the features. Moreover, six machine learning models - SVM and KNN - are deployed for classification. The results of the study demonstrate that the random forest classifier has a diagnostic accuracy of 94%, whereas the SVM has a diagnostic accuracy of 93%.

In Khounraz et al. (2023), A prediction model for the prognosis of COVID-19 patients based on an Iranian data set of COVID-19 patients has been developed. Several approaches were followed in the organization of the research. The process in its entirety entails the gathering of data sets, preparation, choosing features, training, and validation. Based on seven techniques (i.e., logistic regression, gradient-boosted trees, naive bayes, decision trees, support vector machines, generalized linear models, and random forest algorithms), this study evaluated the effectiveness of categorization techniques to forecast COVID-19 mortality. With the best accuracy rates of 86.45 and 84.80%, respectively, the random forest and gradient-boosted tree techniques seemed the most profitable.

4. Backgrounds

4.1. Moth flame optimization (MFO)

The MFO is a recent nature-inspired meta-heuristic paradigm that attempts to imitate the navigation of moths in the night. MFO is a population-based algorithm, and there have been several advanced variants of this method until now (Xu et al., 2019). Moth and flame serve as the foundation for the MFO's mathematical model. While the flames are the moths' best current position, the true search agents are those that move about the search area. The transverse aspect is the algorithm's motivation, as was already stated. The position of each moth relative to a flame is adjusted using mathematics to represent this activity via (1):

$$Moth_i = \mathrm{S}\left(Moth_i, Flame_j\right) \tag{1}$$

Where $Moth_i$ indicates the i^{th} moth, $Flame_j$ refers to the j^{th} flame, and S is the spiral function. The logarithmic spiral for the MFO algorithm is defended by (2):

$$\mathrm{S}\left(Moth_i, Flame_j\right) = D_i \cdot e^{bt} \cdot \cos(2\pi t) + Flame_j \tag{2}$$

Where D_i refers to the distance of the i^{th} moth for the j^{th} flame and is as defined in (3), b is a constant for defining the shape of the logarithmic spiral, and t indicates a random number in [−1, 1]. D is calculated by (3):

$$\mathrm{D_{i=}}\left|Flame_j - Moth_i\right| \tag{3}$$

The t parameter in (2) determines how a moth will navigate around a flame ($t = -1$ is positioned nearest to the flame, otherwise $t = 1$ demonstrates how far). A moth can fly across a flame using the spiral equation, not just in the area among them. As a result, it is possible to ensure the exploration and exploitation of the search space.

Table 1. An overview of recent coronavirus disease 2019 recognition studies.

Research	Year	Overview	Benefits	Problems
Maghdid et al. (2020)	2020	To find COVID-19, chest radiography and CT imaging datasets were both employed. The suggested model incorporates the convolutional neural network and the transfer learning algorithm as its two primary techniques.	Simple to perform The accuracy = 94.1%	The accuracy of the suggested research is not sufficient with limited number of images currently available about COVID-19 cases. Runtime and memory complexities
Sun et al. (2020)	2020	A deep forest model for learning high level representation of features was categorized to classify COVID-19 patients based on using chest CT images. Utilizing the trained forest, an updated feature selection process has been deployed to eliminate feature redundancy and develop classification performance metrics.	The dataset is huge. The accuracy = 91.79%	The previous information in the present employment is used to pick the features. The performance has been promoted using a deep learning method. Runtime and memory complexities
Nazish et al. (2021)	2021	This study suggested a machine learning technique (Support vector machine and logistic regression) to identify COVID-19 and normal patients using chest radiography images.	The accuracy of SVM = 96%	Runtime and memory complexities
Sen et al. (2021)	2021	using meta-heuristic-based wrapper feature selection methods with deep learning to identify COVID-19 illness. The COVID-19 and non-COVID chest CT images have been classified using the final feature set by the SVM classifier.	In order to accurately predict COVID-19 from the participant's chest CT images, features were extracted using CNN and then chosen using the proposed feature selection approach. The accuracy = 90%	Small dataset to detect COVID 19 was used. Time and memory complexities
Zhang et al. (2021)	2021	The proposed framework was proposed for COVID-19 diagnosis using CT. It combines deep convolutional neural network and stochastic pooling.	The proposed framework added batch normalization transform and dropout layers, and also proposed two new blocks (convolution block and fully connected block). The Accuracy = 93.64%	The dataset is somewhat small.
Habib et al. (2022)	2022	By examining chest radiography pictures for the impact of a viral illness, this suggested study explored the potential of machine learning to discriminate between COVID-19, viral pneumonia infected, and normal.	The accuracy of random forest classifier = 94%	Small dataset to detect COVID 19 was used.
Khounraz et al. (2023)	2023	With the use of lab tests and demographic information, data mining approaches have the possibility of being used to predict the outcomes of COVID-19 patients.	The accuracy = 84.8% (with Gradient Boosted trees)	The dataset is somewhat small. The uneven distribution of individual lab tests, which had an impact on data processing

So as to further assert utilization, t is described as random number through $[a,1]$ where a gradually decreases over the period of iteration from -1 to -2. According to (2), each moth is only allowed to move so far towards a flame before it may cause a localized optimal recession. To avoid this, at each iteration, a list of flames should be modified and ordered according to their fitness levels. After then, the moths adjust their locations in relation to their associated flames. An adjustable strategy for the number of flames has been presented since the exploitation of the most viable solutions may suffer from the position-modifying of moths with regard to n various locations in the search space via (4):

$$\text{Flame number} = \text{round}\left(N - l * \frac{N-1}{T}\right) \tag{4}$$

Where l the current number of iteration, N is the maximum number of flames, and T indicates the maximum number of iterations.

4.2. Rough set principle

For the understanding of expert systems that are described by limited and partial information, rough set theory is a variant of set theory (Qinghua et al., 2016). Two describable subsets, known as the lower and higher approximations, roughly depict an indescribable subset. To specify the requirement for features, principles from rough set theory are applied. The indicators of requirement are computed using lower and upper approximation functions. To influence the process of selecting features, these indicators are used as criteria (Deng et al., 2022).

Assume $I = (U, A)$ as an information system, during which U is a nonempty set of finite objects (the universe, or the COVID-19 dataset), as well as A is a non-empty finite collection of features (extracted features set from COVID-19 dataset), this can help to understand the rough set. Therefore, ∀ a ∈ A determines a function f_a: U $\rightarrow V_a$.With any P ⊆ A, there is an associated equivalence relation expressed by (5):

$$\text{IND}(P) = \{(x,y) \in U \times U | \forall a \in P, f_a(x) = f_a(y)\} \tag{5}$$

The partition of U, generated by IND (P), is denoted U/P. The equivalence classes of the P-indiscernibility relation are denoted $[x]_p$. The indiscernibility relation is the mathematical basis of rough set theory.

Let X ⊆ U, the P-lower approximation $\underline{P}$ X, and P-upper approximation $\overline{P}$ X of set X can be defined as in (6) and (7):

$$\underline{P}X = \left\{x \in U \,|\, [x]_p \subseteq X\right\} \tag{6}$$

$$\overline{P}X = \left\{x \in U \,|\, [x]_p \cap X \neq \phi\right\} \tag{7}$$

Let P, Q ⊆ A be equivalence relations over U, then the positive, negative and boundary regions can be defined as in (8),(9), and (10):

$$POS_p(Q) = \cup_{x \in u/Q} \underline{P}X \tag{8}$$

$$NEG_p(Q) = \text{U} - \cup_{x \in u/Q} \overline{P}X \tag{9}$$

$$BND_p(Q) = \cup_{X \in U/Q} \overline{P}X - \cup_{X \in U/Q} \underline{P}X \tag{10}$$

The positive region of the partition U/Q with respect to P ($POS_p(Q)$), is the collection of all U objects that might be grouped into division blocks with certainty. Finding connections between attributes is a critical issue in attribute decline. U/Q by means of P. For P, Q ⊆ A, we state that Q is reliant on P to a degree of k (0 ≤ K ≤ 1) depicted $P \Rightarrow_K Q$, if

$$\text{K} = \gamma_p(Q) = \frac{|\boldsymbol{POS}_p(\text{Q})|}{|U|} \tag{11}$$

If k = 1, Q relies entirely upon P, if 0 < k < 1, Q relies partially (in a degree k) upon P, and if k = 0, Q is independent of P.

In a decision system, an attribute set includes two sets: decision attribute set (i.e. has two values 'yes' patient is infected and 'no' patient is non-infected) D and condition attribute set (i.e. COVID CT scans and Non-COVID CT scans) C, i.e. $A = C \subset D$. The degree of dependency between these two sets, $\gamma_C(D)$, which is known as the quality of approximation of classification, is induced by the decision attributes set. When P is a set of condition attributes and Q is the decision, $\gamma_P(Q)$ is the quality of classification. The goal of attribute reduction is to remove redundant attributes so that the reduced set provides the same quality of classification as the original. A reduce is defined as a subset R of the conditional attribute set C such that $\gamma_R(D) = \gamma_C(D)$. The set of all reduces is defined as in (12):

$$\text{Red} = \{R \subseteq C | \gamma_R(D) = \gamma_C(D), \forall\, B \subseteq R, \gamma_B(D) \neq \gamma_C(D)\} \tag{12}$$

A reduce with a limited cardinality is that which is investigated in a rough set attribute reduction. To identify just one component of the minimal reduce $Red_{min} \subseteq Red$, the later equation is employed:

$$\text{Red} = \{R \in Red | \forall R' \in Red, |R| \leq |R'|\} \tag{13}$$

The intersection of all reduces is called the core, the elements of which are those attributes that cannot be eliminated (i.e. selected feature set). The core is defined as in (14):

$$Core(C) = \cap Red \tag{14}$$

4.3. Support vector machines (SVMs)

SVMs were first used to solve binary classification problems, which are two-class problems (Vapnik, 1999). As shown in Fig. 1, data in binary classification issues is differentiated by a hyperplane defined by a set of support vectors. SVMs can establish nonlinear classification utilizing the kernel method in conjunction with linear classification. SVM uses a number of kernel functions to enable it to find the best solution. The trade-off constant C, and indeed the type of the kernel function, must be determined before SVMs may be deployed. The soft margin is regulated by the parameter C, which determines the effect of each individual support vector. Otherwise, polynomial, sigmoid, and radial basis kernel functions are the most commonly utilized kernel functions (RBF) (Narin, 2020).

5. The suggested COVID-19 patients recognition framework (PRF)

The suggested PRF in the healthcare system will be thoroughly described in this section. The PRF's major aim is to identify COVID-19 instances rapidly and precisely. Early identification of COVID-19 instances enables prompt patient separation and remedy, which reduces the propagation of the disease's infection. The suggested recognition framework, which is shown in Fig. 2, is divided into three stages: (i) the P^2S, (ii) the FS^2, and (iii) the CS. The major aim of P^2S is to extract a collection of features from a CT image using GLCM and then remove any extraneous features using the EMFO approach. This scheme only records the most valuable data, allowing the subsequent stage, known as CS, to recognize COVID-19 instances promptly and precisely. SVM classifiers can be used in CS to quickly and accurately classify individuals who are diseased by using the selected features from CS.

5.1. Feature selection stage (FS2)

To achieve FS^2, the EMFO approach has been used by assessing the ideal collection of features and using both the rough set and the SVM classifier as fitness function for the MFO in order to reach the best accuracy. The MFO algorithm was used in this study for the following reasons: (i) According to the original publication introducing the MFO (Xu et al., 2019), the MFO algorithm offers advantages over comparable algorithms like particle swarm optimization (Kennedy and Eberhart, 1997), genetic algorithm (Kabir et al., 2011), and binary grey wolf optimizer (Mirjalili et al., 2014) in the scope of optimization issues. (ii) As long as moths adjust their positions as per flames, which are the most promising alternatives, the MFO convergence is maintained. (iii) Due to its straightforward, adaptable, and simple implementation techniques, MFO

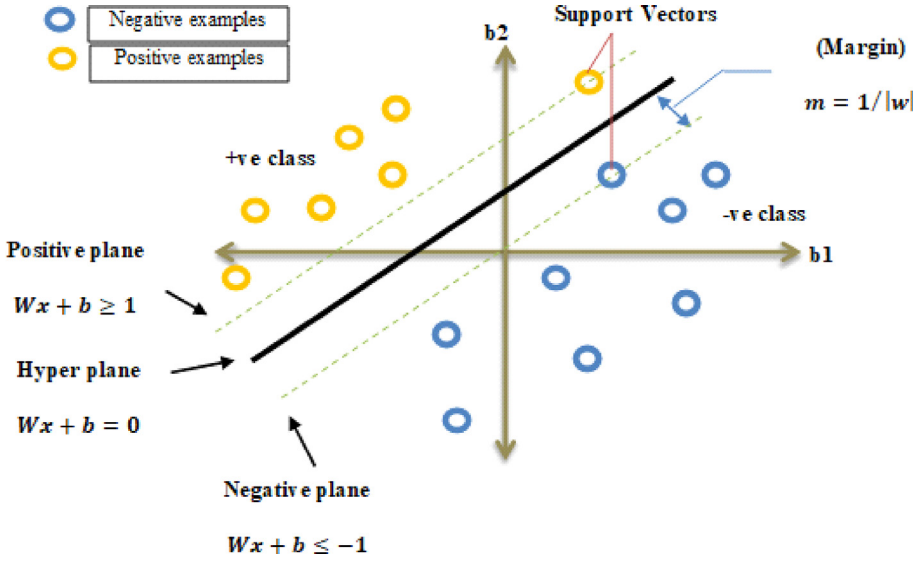

Fig. 1. Distinguishable hyper plane among two datasets.

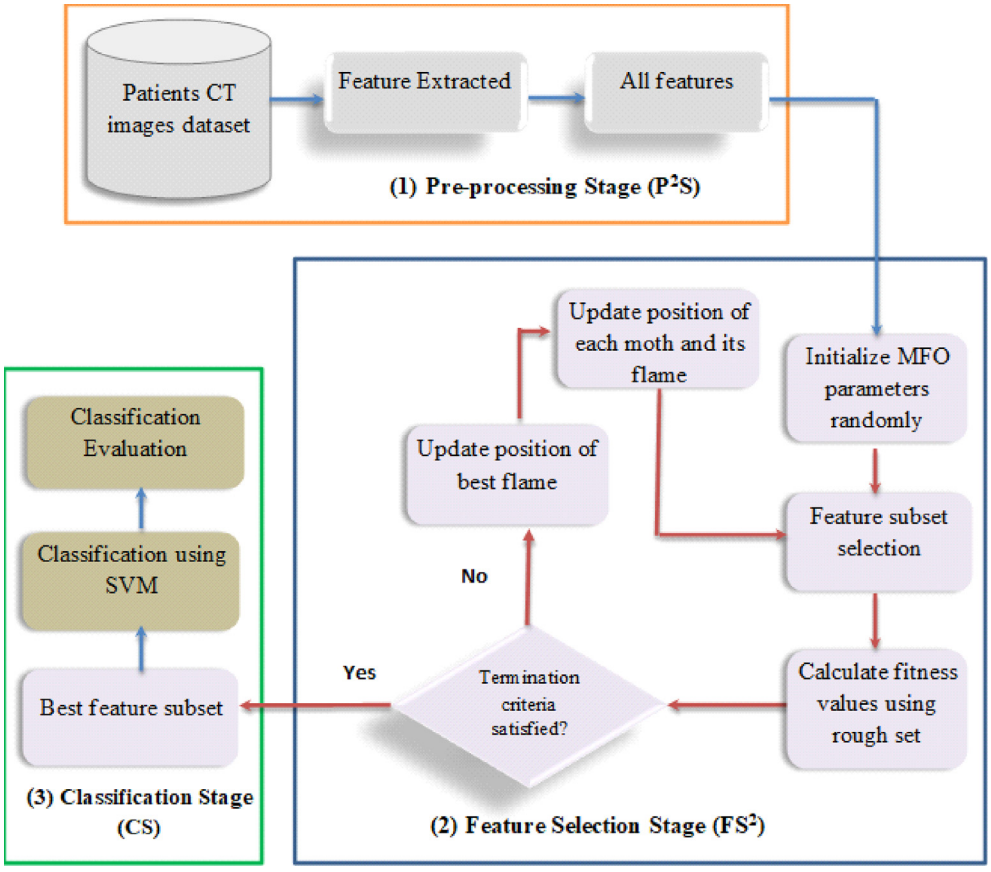

Fig. 2. The suggested coronavirus disease 2019 Patients Recognition Framework (PRF).

can be used to address a variety of issues (Jangir et al., 2016). The overall proposed EMFO feature selection approach is described in Algorithm 1.

In the EMFO approach, the solution space indicates all possible selections of features (a subset of informative features). Each moth position indicates a

Algorithm 1. The suggested EMFO

```
Initiate MFO parameters.
    For (i = 1: number of flame) do
        Initiate the population of solutions via Eq. (15)
        Examine every moth's fitness via Eq. (16)
    End for
        Arrange the first moth population.
        Adjust the location of the best flame found.
while (Iter < Max_Iter and Global_FlamFit < MaxFit) do

    Adjust flame number via Eq. (4)
    Arrange moths per respective fitness values and allocate the highest one.
    Adapt flame locations to the ordered moths.
    Reduce a's value from _1 to 2.
For (i = 1: number of flame) do
        Adjusted region-specific location of each moth (feature set) [1, N/3] via Eq. (15)
        Adjust each flame's location in accordance with the best moth
End for
 Examine each moth's fitness via Eq. (16)
 iter = iter + 1;
End while
 Establish the optimal flame location.
```

Algorithm Parameters	
γ_R	classification quality of condition attribute set R
$\|C\|$	Total number of features.
D	Decision
β	Parameter corresponding to the subset length
α	Parameter corresponding to the importance of classification quality
$\|R\|$	length of elected attribute subset

binary selection of feature sets of length N, where N is the total number of features (i.e., equal to the number of features in the COVID-19 dataset). Every bit represents a feature, and a value of '1' indicates that the feature is selected, while a value of '0' indicates that it is not. Each position is a subset of features. For each moth, the periodicity of a position update is represented as a positive integer with a range of 1 to max-update. It refers to the number of bits (features) that should be altered for the moth at a specific time.

The widest range of location updates restricts a moth's ability to perform global exploration. The maximum for each moth's position update was found to be (1/3) *N after extensive testing. Studies like (Wang et al., 2007), among others, have demonstrated that this upper limit results in efficient results.

When the population is randomly initialized, a feature subset (solution) should be produced randomly by (15).

$$X_{ij} = \begin{cases} 1, rand(\,) > 1 \\ 0, x \geq otherwise \end{cases} \tag{15}$$

Where, i $\in\{1,2,\ldots\ldots.PN\}$ and j $\in\{1,2,\ldots\ldots.FN\}$ where PN is population size and FN is number of feature.

The fitness function is a measure to determine the goodness or quality of a single solution (a feature subset) in a population. As a result of the evaluation of quality search, the fitness value is calculated for each agent at each individual iteration. In EMFO, classification accuracy was employed as a fitness function, and an SVM classifier was utilized to evaluate the performance of each solution (i.e., to select the most effective features for COVID-19 diagnosis). The average of the 10-fold cross-validation procedure served as the foundation for the classification accuracy that was attained. Given that we must evaluate the classification accuracy and the size of the feature portion, the fitness function is calculated using (16):

$$\text{Fitness Function} = \alpha * \gamma_R(D) + \beta * \frac{|C| - |R|}{|C|} \tag{16}$$

where:

(1) $\gamma_R(D)$: is the condition feature set R's classification accuracy in relation to the decision D.
(2) $|R|$: the length of the chosen feature portion.
(3) $|C|$: the total number of features.
(4) α and β are factors that, in turn, signify the importance of classification accuracy and portion length, $\alpha \in [0, 1]$; and β = 1-α. We believe that classification accuracy is more important than portion size and set α = 0.9; β = 0.1.

5.2. Classification stage (CS)

During the CS, the SVM was also used to differentiate COVID-19 patients from non-COVID-19 patients. In this stage, after classifying the data, the framework is trained and validated. A confusion matrix is produced as a graphic form of performance. Each row refers to the instances in its real class, whereas each column refers to the instances in a predicted class. Based on this matrix, the recall, error, precision, accuracy, and F-measure are calculated to evaluate the classifier.

5.3. Complexity of the proposed PRF

In this sub-section, we discuss the computational complexity of the PRF. In general, its complexity depends on the complexity of the MFO. The MFO algorithm's computational complexity is reliant on the number of moths, number of variables, maximum number of iterations, and the flames' individual iteration sorting mechanism (Mirjalili, 2015). Since MFO adopts the Quicksort method, the computational complexity is O (n^2) in the worst case. Hence, the overall computational complexity is defined as follows:

$$\begin{aligned} O\,(\text{MFO}) = \big(O\,\big(t\,\big(O\,(\text{Quick sort}) \\ + O\,(\text{position of update})\big)\big)\big) \end{aligned} \tag{17}$$

$$O\,(\text{MFO}) = O\,\big(t\big(n^2 + n \times d\big)\big) = O\,\big(tn^2 + tnd\big) \tag{18}$$

where, n is the number of moths, t is the maximum number of iterations, and d is the number of variables.

6. Experimental results

Through this segment, the suggested PRF will be evaluated. PRF is implemented through the following steps: (i) the feature selection method using the EMFO approach; and (ii) the classification method using SVM. Before applying the EMFO approach and afterwards the SVM method, the feature extraction procedure will first be carried out using the GLCM method to extract a subset of features from CT images. Second, the suggested EMFO approach will be used to choose the most important features from those extracted. The suggested PRF will then be used to accurately identify COVID-19 cases with the least amount of time spent. Our implementation is based on the COVID-CT dataset (Zhao et al., 2020), a publicly available dataset called COVID-CT that includes 463 non-COVID-19 CT findings and 349 COVID-19 CT depictions from 216 individuals. Confirming the

usefulness of this dataset is a prominent radiologist who successfully identified and managed COVID-19 patients prior to the epidemic occurring. Among the patients with a positive label, 169 have age and 137 have gender information. There are more male patients than female ones (86 vs. 51). The COVID-CT dataset (Zhao et al., 2020) is split into two sets: training and testing. For model learning, the training set is utilized. The suggested model's suitability is then evaluated using the testing set. In Table 2, the exact values for the feature selection approach's applicable parameters are shown.

6.1. Evaluation metrics

The following experiments will calculate four assessment parameters: accuracy, error, precision, sensitivity, and f-measure. As well, micro-average will be measured in connection with precision and recall as additional parameters, further clarifying the application findings. A confusion matrix is employed to determine the values of various measurements, as shown in Table 3. Therefore, various equations are employed to describe the confusion matrix, as shown in Table 4 (Ali et al., 2020; Ali, 2021).

Table 2. The specified variables with the appropriate utilized values.

Variables	utilized value
No. of population	30
No. of generation	50
Velocity	$1 \sim (1/3)*N$
Wight	$1.4 \sim 0.4$
α	0.9
β	0.1

Table 3. Confusion matrix.

	Anticipated as positive	Anticipated as negative
Truly Positive	True positive (TP)	False negative (FN)
Truly Negative	False positive (FP)	True negative (TN)

6.2. Testing the proposed EMFO feature selection approach

In this segment, EMFO will be implemented and tested against other contemporary selection approaches via Table 5 to represent its efficiency in selecting the best features in the COVID-19 CT dataset. The SVM classifier will be deployed as the basic classifier to test these approaches and show the wellness of EMFO compared with other approaches. Tables 6–8 show the results. The accuracy, error, precision, and recall measurements in Table 6 show how well the EMFO approach performs in comparison with modern approaches. The deployment findings in Table 7 (macro-average precision, macro-average recall, micro-average precision, and micro-average recall) then demonstrate how well the EMFO approach performs in comparison with other modern approaches. Finally, Table 8's results, which include run time and F-measure, show how the EMFO approach compares with several other contemporary techniques.

Table 6 demonstrates that all techniques perform better when there are more patients in the training data. Maximum patient training (e.g., 498 patients) leads to high precision, recall, and accuracy, as well as the lowest error. Because as the number of training cases grows, so will the data that is collected and the number of classification rules. As a result of the classifiers' improved training, classification accuracy is also improved. Additionally, it is stated that the intended EMFO approach presents the best performance compared with other recent approaches. As a result, while FP and FN are decreased, TP and TN are both increased. This increases the proposed selection approach's accuracy, precision, and recall while lowering its error.

Otherwise, SDS provides the weakest achievement because it removed a useful feature, and the SVM classifier was subsequently trained on individuals utilizing the least reliable set of features.

Table 4. Confusion matrix equations.

Measure	Equation	Description
Precision (P)	TP/(TP + FP)	The proportion of correct positive predictions.
Recall/sensitivity (R)	TP/(TP + FN)	The proportion of instances with positive labels that were also projected to be positive.
Accuracy (A)	(TP + TN)/(TP + TN + FP + FN)	The proportion of predictions that are correct.
Error (E)	1-Accuracy	The proportion of predictions that are incorrect.
Macro-average	$\sum_{i=1}^{c} P_i/c$ ' for Precision' $\sum_{i=1}^{c} R_i/c$ ' for Recall'	The average of the precision and recall of the system on different c classes.
Micro-average	(TP1+TP2)/(TP1+TP2+FP1+FP2) 'for precision' (TP1+TP2)/(TP1+TP2+FN1+FN2) 'for Recall'	The summation up to the individual true positives, false positives, and false negatives of the system for different classes and the apply them to get the statistics.
F-measure	2*PR/(P + R)	The weighted harmonic mean of Precision and Recall.

Table 5. Latest methods of feature selection utilized in assessment.

Methods	Explanation
SDS (Shanth and Rajkumar, 2021)	The best feature subsets have been categorized based on a proposed feature selection method. Initially, each agent is tasked with integrating the feature subset from their respective search spaces. For both the training and testing groups, each agent now uses an independent, random division of the dataset.
BRSA (Krishanthi et al., 2023)	A novel efficient feature selection technique is presented that effectively selects and minimizes the dimensionality of a gene activity dataset by combining the LASSO regression method with the binary reptile search algorithm.
MQMPA (Torse et al., 2023)	A modified quantum-based marine predator's algorithm (MQMPA) feature selection method has been presented. The proposed MQMPA has been employed on COVID19 CT images dataset.
LNNLS-KH (Li et al., 2021)	It has been proposed for feature selection in network intrusion detection. The number of selected features and classification accuracy are introduced into the fitness evaluation function of the LNNLS-KH algorithm, and the physical diffusion motion of the krill individuals is transformed by a nonlinear method. Thus, the linear nearest neighbor lasso step optimization is performed on the updated krill herd position in order to derive the global optimal solution.
CGAFS (Rostami et al., 2021)	It proposes a genetic algorithm based on community detection that functions in three steps. The feature similarities are calculated in the first step. The features are classified by community detection algorithms into clusters throughout the second step. In the third step, features are selected by a genetic algorithm with a new community-based repair operation.

EMFO provides around 0.92 precision for training patients (498), compared with 0.8 for SDS. SDS and EMFO recall values are 0.78 and 0.91, respectively, at 498 training patients. BRSA's accuracy is 0.81, whereas EMFO's is 0.94 as the number of training patients approaches 498, demonstrating that SDS has a higher mistake rate than EMFO.

The findings in Table 7 show that EMFO has the best macro-average precision with a value of 0.9, while SDS has the worst value at training patients,

Table 6. Comparisons between enhanced moth flame optimization and the recent feature selection methods using accuracy, error, precision, and recall.

No. of training patients	Accuracy (%)						Error (%)					
	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO
70	69	71	74	76	78	80	31	29	26	24	22	20
140	71	73	76	78	80	83	29	27	24	22	20	17
210	72	75	78	80	83	86	28	25	22	20	17	14
280	75	77	80	82	85	88	25	23	20	18	15	12
350	77	79	82	84	87	90	23	21	18	16	13	10
420	79	81	84	87	89	92	21	19	16	13	11	8
498	81	83	86	89	91	94	19	17	14	11	9	6
AVG.	74.85	77	80	82.28	84.7	87.57	25.14	23	20	17.71	15.28	12.42
STD.	4.41	4.32	4.32	4.71	4.71	4.96	4.41	4.32	4.32	4.71	4.71	4.96
Median	75	77	80	82	85	88	25	23	20	18	15	12
No. of training patients	Precision (%)						Recall (%)					
	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO
70	68	71	73	74	75	78	67	69	70	71	72	76
140	70	72	75	76	77	80	69	70	73	74	75	78
210	71	73	77	78	79	82	70	72	74	75	77	80
280	73	75	79	80	81	85	72	75	77	78	79	83
350	76	78	80	82	83	87	74	77	79	80	82	86
420	78	80	82	84	86	90	76	79	81	82	84	89
498	80	82	84	86	88	92	78	81	83	84	86	91
AVG.	73.71	75.85	78.57	80	81.28	84.85	72.28	74.71	76.71	77.71	79.28	83.28
STD.	4.42	4.42	3.86	4.32	4.715	5.17	3.94	4.57	4.64	4.64	5.02	5.64
Median	73	75	79	80	81	85	72	75	77	78	79	83

Table 7. Comparisons between enhanced moth flame optimization and the recent feature selection methods using micro-average precision, micro-average recall, macro-Average precision, and macro-average recall.

No. of training patients	Micro-average precision (%)						Micro-average recall (%)					
	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO
70	70	71	73	75	76	80	68	70	71	74	75	78
140	71	73	75	76	78	82	69	71	73	75	77	81
210	73	75	76	78	80	84	70	72	75	77	79	82
280	75	76	78	80	81	86	72	74	77	79	80	85
350	76	78	80	81	83	88	74	76	79	80	82	87
420	78	80	81	83	85	90	76	78	80	82	84	89
498	80	81	83	85	87	91	77	79	82	84	86	90
AVG.	74.71	76.28	78	79.71	81.42	85.85	72.28	74.28	76.71	78.71	80.42	84.57
STD.	3.638	3.638	3.559	3.636	3.866	4.099	3.49	3.49	3.94	3.63	3.86	4.42
Median	75	76	78	80	81	86	72	74	77	79	80	85
No. of training patients	Macro-average Precision (%)						Macro-average Recall (%)					
	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO
70	70	70	74	77	78	82	67	69	72	76	77	81
140	72	71	76	79	80	84	68	71	74	77	78	82
210	73	73	77	80	81	85	70	73	76	78	81	84
280	75	75	78	81	83	87	71	74	77	80	82	85
350	76	77	80	82	84	88	73	76	79	81	83	86
420	78	79	81	83	85	89	75	77	80	82	84	88
498	79	80	82	84	86	90	76	78	81	83	85	89
AVG.	74.71	75	78.28	80.85	82.428	86.428	71.42	74	77	79.57	81.428	85
STD.	3.25	3.87	2.87	2.41	2.87	2.878	3.408	3.265	3.265	3.265	2.99	2.94
Median	75	75	78	81	83	87	71	74	77	80	82	85

Table 8. Comparisons between enhanced moth flame optimization and the recent feature selection methods using Run-time and F1-measure.

No. of training patients	Run-time (sec.)						F1-measure (%)					
	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS	EMFO
70	6	3	5	4	5	2	67.496	69.985	71.468	72.469	73.469	76.987
140	8	5	7	6	7	4	69.496	70.985	73.986	74.986	75.986	78.987
210	10	6	8	7	8	5	70.496	72.496	75.470	76.470	77.987	80.987
280	12	8	10	8	9	6	72.496	75	77.987	78.987	79.987	83.988
350	13	9	11	9	10	7	74.986	77.496	79.496	80.987	82.497	86.497
420	14	11	12	11	11	8	76.987	79.496	81.496	82.988	84.988	89.497
498	15	12	14	13	12	9	78.987	81.496	83.988	84.988	86.988	91.497
AVG.	11.14	7.7	9.57	8.28	8.86	5.86	72.99	75.279	77.629	78.839	80.272	84.063
STD.	3.28	3.25	3.1	3	2.4	2.4	4.178	4.390	4.260	4.485	4.869	5.416
Median	12	8	10	8	9	6	72.496	75	77.987	78.987	79.987	83.988

reaching 498 with a value of 0.79. The macro-average recall of training patients is equivalent to 498, with EMFO having the greatest value of 0.89 and SDS having the lowest value of 0.76. EMFO introduces a roughly 0.91 micro-average precision value while training patients to equal 498, compared with 0.80 for SDS. Additionally, the micro-average recall value for the SDS, community genetic algorithm for feature selection (CGAFSS), and EMFO at training patients reaches 498, which is 0.77, 0.86, and 0.9 separately. Table 10 shows that when training patients equals 498, the best F-measure value is 0.67496 and the worst value is 0.7898, both of which are related to SDS. Furthermore, EMFO implementation takes less time than SDS when it comes to training sufferers, who total 498 with run times of 9 and 15 s, respectively.

6.3. Testing the suggested patients recognition framework (PRF)

The main goal of this experiment is to test the whole suggested PRF. To demonstrate the efficacy of our suggested framework, it is contrasted with a few recently employed COVID-19 recognition studies, as shown in Table 1. Recent research studies are (Maghdid et al., 2020; Sun et al., 2020; Nazish et al., 2021; Sen et al., 2021; Zhang et al., 2021; Habib et al., 2022; Khounraz et al., 2023). As a result, the EMFO approach is employed for feature selection, and SVM is used for classification in our PRF, which takes advantage of all capabilities presented. PRF's error decreases as its accuracy, precision, recall, and F-measure improve. Moreover, as illustrated in Table 10, the proposed PRF demonstrates the best

performance in terms of macro-average precision, micro-average precision, macro-average recall, micro-average recall, and run-time. This demonstrates the value of EMFO and SVM, the two key components of the suggested PRF, and shows their ability to function well together. PRF demonstrated that it is quicker than other approaches as well.

Table 9 shows that Fariba K. et al. (Khounraz et al., 2023) provide an accuracy of 0.80, whereas the PRF at training patients = 498 is 0.98. Based on the most important features selected by EMFO, SVM accurately identifies infectious patients in the shortest amount of time. As a result, while Fariba K. et al. (Khounraz et al., 2023) introduce the maximum

Table 9. Comparisons between patients recognition framework and the existing coronavirus disease 2019 recognition methods using accuracy, error, precision, recall and f1-measure.

No. of training patients	Accuracy (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	67	71	73	75	77	79	81	85
140	69	73	75	77	80	81	83	89
210	71	75	77	80	83	85	86	91
280	73	78	80	83	85	87	88	92
350	76	82	82	85	88	89	90	94
420	78	85	87	88	90	91	93	96
498	80	87	89	91	93	94	96	98
AVG.	73.42857	78.71	80.428	82.71	85.14	86.57	88.14	92.14
STD.	4.790864	6.129	5.99	5.79	5.639	5.349	5.33	4.37
Median	73	78	80	83	85	87	88	92
No. of training patients	Error (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	33	29	27	25	23	21	19	15
140	31	27	25	23	20	19	17	11
210	29	25	23	20	17	15	14	9
280	27	22	20	17	15	13	12	8
350	24	18	18	15	12	11	10	6
420	22	15	13	12	10	9	7	4
498	20	13	11	9	7	6	4	2
AVG.	26.57143	21.28	19.57	17.28	14.85	13.428	11.85	7.85
STD.	4.790864	6.12	5.9	5.79	5.63	5.34	5.33	4.37
Median	27	22	20	17	15	13	12	8
No. of training patients	Precision (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	65	67	66	70	71	73	77	80
140	66	68	68	71	74	75	79	82
210	68	70	72	74	76	78	81	85
280	70	71	74	76	79	81	83	87
350	71	73	76	78	81	83	85	89
420	73	75	78	80	83	85	88	91
498	75	77	80	82	85	87	90	93
AVG.	69.714	71.57	73.42	75.85	78.42	80.28	83.28	86.71
STD.	3.63	3.64	5.12	4.48	5.02	5.18	4.715	4.715
Median	70	71	74	76	79	81	83	87
No. of training patients	Recall (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	61	64	65	69	71	74	76	79
140	63	66	66	71	74	76	78	81
210	66	68	68	73	76	78	80	83
280	68	70	70	76	78	80	83	86
350	70	72	73	78	80	82	85	88

(*continued on next page*)

Table 9. (continued)

No. of training patients	Accuracy (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
420	72	74	76	80	82	84	87	90
498	74	76	78	81	84	86	89	92
AVG.	67.71	70	70.85	75.42	77.85	80	82.57	85.57
STD.	4.71	4.32	4.98	4.57	4.56	4.32	4.79	4.79
Median	68	70	70	76	78	80	83	86
No. of training patients	F1-measure (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	62.9365	65.4656	65.4918	69.4964	71	73.4966	76.496	79.496
140	64.4651	66.9851	66.98507	71	74	75.4966	78.496	81.496
210	66.9851	68.9855	69.94286	73.4966	76	78	80.496	83.988
280	68.9855	70.4965	71.9444	76	78.4968	80.496	83	86.497
350	70.4965	72.4966	74.4698	78	80.4968	82.496	85	88.49
420	72.4966	74.4966	76.98701	80	82.496	84.496	87.497	90.49
498	74.4966	76.4967	78.98734	81.496	84.497	86.497	89.497	92.49
AVG.	68.6946	70.7747	72.1161	75.64142	78.1411	80.140	82.926	86.138
STD.	4.19159	3.98717	5.0229	4.525	4.7919	4.749	4.747	4.75
Median	68.9855	70.4965	71.9444	76	78.4968	80.4968	83	86.497

error value of 0.2, PRF introduces the minimum value of 0.02. While it is 0.75 for Fariba K. et al. (Khounraz et al., 2023) at training patients (i.e., 498), PRF offers about 0.93 precision. The PRF recall value is 0.92; however, when the training patients are equivalent to 498, it is 0.74 for Fariba K. et al. (Khounraz et al., 2023). When training patients = 498, PRF introduces about 0.9249 F-measure, while it is 0.7449 for Fariba K. et al. (Khounraz et al., 2023). Consequently, Table 9 illustrates that PRF is much better than recent studies in Maghdid et al., 2020; Sun et al., 2020; Sen et al., 2021; Habib et al., 2022; Khounraz et al., 2023.

Table 10's findings reveal that the macro-average PRF precision is 0.91, compared with 0.76 for Fariba K. et al. (Khounraz et al., 2023) when training patients equivalent to 498. Moreover, when training patients equals 498, PRF introduces roughly 0.88 macro-average recall compared with 0.72 for Fariba K. et al. (Khounraz et al., 2023). While Fariba K. et al.

Table 10. Comparisons between patients recognition framework and the existing coronavirus disease 2019 recognition methods using micro-average precision, micro-average recall, macro-average precision, macro-average recall, and run-time.

No. of training patients	Micro-average precision (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	61	62	63	70	71	72	74	78
140	64	65	66	72	73	74	76	80
210	67	68	69	74	75	76	78	82
280	70	71	71	76	77	78	80	84
350	72	73	74	78	80	82	82	85
420	74	75	77	80	81	84	85	87
498	75	77	79	81	83	85	87	89
AVG.	69	70.1	71.28	75.85	77.14	78.71	80.28	83.57
STD.	5.228	5.42	5.79	4.09	4.41	5.05	4.71	3.86
Median	70	71	71	76	77	78	80	84
No. of training patients	Micro-average Recall (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	61	62	62	63	68	70	72	72
140	63	64	64	67	70	72	74	74
210	65	66	66	70	72	74	76	76
280	66	68	68	72	74	76	78	78
350	68	70	71	74	76	78	80	81

(*continued on next page*)

Table 10. (continued)

No. of training patients	Micro-average precision (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
420	70	72	73	76	78	80	82	83
498	71	74	76	78	80	82	84	86
AVG.	66.285	68	68.57	71.42	74	76	78	78.57
STD.	3.634	4.32	5.02	5.22	4.32	4.3	4.3	5.03
Median	66	68	68	72	74	76	78	78
No. of training patients	Macro-average Precision (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	64	67	68	70	74	75	78	80
140	66	68	70	71	76	77	79	81
210	68	70	71	73	77	79	81	83
280	70	72	73	75	79	81	83	85
350	72	74	75	77	81	83	85	87
420	75	76	77	80	83	85	87	89
498	76	78	80	83	85	87	89	91
AVG.	70.14	72.14	73.42	75.57	79.28	81	83.14	85.14
STD.	4.48	4.09	4.19	4.75	3.94	4.32	4.09	4.09
Median	70	72	73	75	79	81	83	85
No. of training patients	Macro-average Recall (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	60	61	63	68	70	71	74	75
140	62	62	66	70	72	73	76	77
210	64	65	68	72	74	75	78	80
280	66	67	71	74	76	77	80	82
350	68	70	73	76	78	80	82	84
420	70	72	75	78	80	82	84	86
498	72	75	77	80	82	84	86	88
AVG.	66	67.42	70.42	74	76	77.42	80	81.71
STD.	4.32	5.19	5.02	4.32	4.32	4.79	4.32	4.715
Median	66	67	71	74	76	77	80	82
No. of training patients	Run-time (sec.) (%)							
	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)	PRF
70	20	24	25	18	17	27	14	4
140	21	25	26	19	18	28	15	4.5
210	22	29	30	20	19	29	16	5
280	23	30	31	21	20	30	17	6
350	24	32	34	23	21	32	18	7
420	25	35	36	26	25	36	18.5	8
498	26	36	37	29	28	39	19	9
AVG.	23	30.14	31.3	22.28	21.14	31.57	16.78	6.2
STD.	2.16	4.6	4.68	3.9	3.9	4.4	1.8	1.8
Median	23	30	31	21	20	30	17	6

(Khounraz et al., 2023) provide the lowest value, equal to 0.75, at training patients equal to 498, PRF introduces the maximum micro-average precision with a value that rises to 0.89. In addition, the PRF's micro-average recall value is 0.86, while it is 0.71 for Fariba K. et al. (Khounraz et al., 2023) in training patients with a 498-patient sample. Table 12 shows that PRF is significantly superior to recent studies in (Maghdid et al., 2020; Sun et al., 2020; Nazish et al., 2021; Sen et al., 2021; Zhang et al., 2021; Habib et al., 2022; Khounraz et al., 2023). In contrast to its rivals, PRF offers quick classification, as seen in Table 10. This occurred as a result of other rivals using the deep learning aspect. Deep learning is considered to be computationally intensive and to demand a significant amount of memory space and computational capabilities. Additionally, it has a significant time drawback. However, PRF is easier to use, more

adaptable, and better suited to deal with issues arising from erroneous data. Furthermore, it depends on an ideal augmented feature selection methodology that chooses only those functional features. The PRF's classification performance during testing (recognition) is unaffected by feature selection because it only occurs once. This precise feature selection method reduces the variance of the used feature set, which in turn cuts down on the amount of time SVM needs to diagnose. To avoid overburdening the healthcare system, PRF represents a quick and precise judgement approach for identifying COVID-19 patients. The following factors can be inferred as to why the suggested PRF operates better than current recognition methods:

(1) The suggested PRF relies on an effective, improved feature selection approach that efficiently chooses the optimal set of features to differentiate the situation at hand.
(2) Based on the optimal features chosen by EMFO, SVM can accurately identify infected individuals with the least amount of time delay as compared with the latest approaches.

6.4. Statistical test

An ANOVA test (one-way ANOVA) is applied to measure the statistical differences between the proposed PRF and other methods that are used for comparison in terms of accuracy and run-time. The hypothesis testing can be formulated here in terms of two hypotheses; the null hypothesis (H0: $\mu_{A1} = \mu_{B1} = \mu_{C1} = \mu_{D1} = \mu_{E1} = \mu_{F1} = \mu_{G1} = \mu_{Y1}$), where A1: PRF, B1 (Khounraz et al., 2023):, C1 (Sen et al., 2021):, D1 (Sun et al., 2020):, E1 (Zhang et al., 2021):, F1 (Habib et al., 2022):, G1 (Maghdid et al., 2020):, and Y1 (Nazish et al., 2021):, and the alternate hypothesis (H1: Means are not all equal). The ANOVA test results are shown in Table 11. Based on these test results, the alternate hypothesis H1 is accepted.

A one-tailed T-test at 0.05 (significance level) is performed to measure the statistical differences between the proposed PRF and other techniques that are used for comparison. The hypothesis testing can be formulated here in terms of two hypotheses; the null hypothesis (H0: $\mu_{A1} = \mu_{B1} = \mu_{C1} = \mu_{D1} = \mu_{E1} = \mu_{F1} = \mu_{G1} = \mu_{Y1}$), where A1: PRF, B1 (Khounraz et al., 2023):, C1 (Sen et al., 2021):, D1 (Sun et al., 2020):, E1 (Zhang et al., 2021):, F1 (Habib et al., 2022):, G1 (Maghdid et al., 2020):, and Y1 (Nazish et al., 2021):, and the alternate hypothesis (H1: Means are not all equal). The results in Table 12 show that the p values are less than 0.05 which indicates that there is a statistically significant difference between groups. Thus, the alternate hypothesis H1 is accepted.

An ANOVA test is applied to measure the statistical differences between the proposed EMFO and other strategies that are used for comparison in terms of accuracy and run-time. The hypothesis testing can be formulated here in terms of two hypotheses; the null hypothesis (H_0: $\mu_{A1} = \mu_{B1} = \mu_{C1} = \mu_{D1} = \mu_{E1} = \mu_{F1}$), where A1: EMFO, B1: SDS, C1: BRSA, D1: MQMPA, E1: LNNLS-KH, and F1: CGAFSS, and the alternate hypothesis (H1: Means are not all equal). The ANOVA test results are shown in Table 13. Based on this test results, the alternate hypothesis H1 is accepted.

Furthermore, a one-tailed T-test at 0.05 (significance level) is performed to measure the statistical differences between the proposed EMFO and other techniques that are used for comparison. The hypothesis testing can be formulated here in terms of two hypotheses; the null hypothesis (H0: $\mu_{A1} = \mu_{B1} = \mu_{C1} = \mu_{D1} = \mu_{E1} = \mu_{F1}$), where A1: EMFO, B1: SDS, C1: BRSA, D1: MQMPA, E1: LNNLS-KH, and F1: CGAFSS, and the alternate hypothesis (H1: Means are not all equal). The results

Table 11. Generated P values of the T-test between patients recognition framework and other relevant coronavirus disease 2019 recognition methods.

Technique/ metric	Khounraz et al. (2023)	Sen et al. (2021)	Sun et al. (2020)	Zhang et al. (2021)	Habib et al. (2022)	Maghdid et al. (2020)	Nazish et al. (2021)
Accuracy	1.68664E-09	1.56204E-06	2.70645E-06	3.3482E-06	6.11311E-06	1.24472E-05	8.88308E-05
Precision	9.56319E-09	2.58753E-08	2.40771E-10	1.74666E-10	4.07257E-09	3.2019E-08	1.33839E-06
Recall	8.83845E-12	1.60566E-10	1.79827E-09	2.62644E-10	6.24532E-09	7.51651E-08	–
F-measure	1.45003E-10	2.6486E-09	6.85725E-11	4.2477E-11	2.17034E-10	1.95422E-20	3.0289E-08
Run time (sec.)	1.6285E-11	2.4768E-07	2.2542E-07	5.70099E-07	9.7734–07	1.1834E-07	5.8199E-10

Table 12. A one-way analysis of variance (ANOVA) test results for patients recognition framework in terms of run-time.

Source of variation	SS	df	MS	F	*P* value	F crit
Between Groups	3620.196	7	517.1709	39.023	1.06E-17	2.207436
Within Groups	636.1429	48	13.25298	–	–	–
Total	4256.339	55		–	–	–

Table 13. A one-way analysis of variance (ANOVA) test results for enhanced moth flame optimization in terms of accuracy and run-time.

ANOVA table	Accuracy			Run-Time		
Source of variations	Between groups	Within groups	Total	Between groups	Within groups	Total
SS	793.3571	755.4286	1548.786	111.1429	311.1429	422.2857
df	5	36	41	5	36	41
MS	158.6714	20.98413	–	22.22857	8.642858	–
F	7.561498	–	–	2.5719	–	–
P-value	6.12E-05	–	–	0.043408	–	–
F crit	2.477169	–	–	2.477169	–	–

Table 14. Generated P values of the T-test between enhanced moth flame optimization and other relevant feature selection techniques.

Technique/metric	SDS	BRSA	MQMPA	LNNLS-KH	CGAFSS
Accuracy	1.703E-08	1.651E-08	1.209E-07	8.043E-07	5.071E-07
Precision	2.687E-08	4.230E-07	9.91264E-06	3.684E-06	1.05205E-06
Recall	1.419E-06	5.071E-07	2.43E-06	6.380E-06	6.491E-06
F-measure	2.289E-07	3.184E-07	4.528E-06	4.314E-06	2.100E-06
Run time (sec.)	3.1098E-06	0.00078	6.380E-06	9.07E-05	–

in Table 14 show that the *P* values are less than 0.05 which indicates that there is a statistically significant difference between groups. Thus, the alternate hypothesis H1 is accepted.

7. Discussion

The investigations conducted in this research rely on two evaluations to test the performance and accuracy of the suggested framework for COVID-19 recognition. Actually, MATLAB 2018a was used on a laptop running Intel (R) Core (TM) i5-10210U and @2.11 G with 16.0 GB of RAM to do the simulation on a single platform. This laptop also comes with the Windows 10 (64-bit) operating system.

As a baseline classifier, the suggested EMFO will be assessed with the SVM classifier. Numerous feature selection approaches are contrasted with the proposed EMFO in order to demonstrate its efficacy. The utmost contemporary methods for evaluating feature selection include SDS (Shanth and Rajkumar, 2021), BRSA (Krishanthi et al., 2023), MQMPA (Torse et al., 2023), LNNLS-KH (Li et al., 2021), and CGAFS (Rostami et al., 2021). Results are depicted in Tables 8–10.

The accuracy, precision, recall, F1-measure, macro-average precision, macro-average recall, micro-average precision, micro-average recall, and run-time for EMFO are 94, 92, 91, 91.49, 90, 89, 91, 90%, and 9 s, respectively, at (training patient = 498) as described in Tables 6–8. Hence, EMFO can offer a quick and effective way to select features. As a result, EMFO outperforms SDS, BRSA, MQMPA, LNNLS-KH, and CGAFS. As TP and TN are both increasing, FP and FN are both decreasing. This improves the assessment parameters of the suggested selection method. Furthermore, when compared with other methodologies, the suggested EMFO has a low standard deviation, demonstrating the reliability and resilience of the suggested approach. Moreover, EMFO provides better performance in terms of average and median. SDS, as opposed to that, yields the lowest results. This occurred because the SDS approach removed a useful feature, and the SVM classifier was subsequently trained on individuals using the least useful set of features.

Second, the proposed PRF will be evaluated. Consequently, the SVM classifier is utilized for classification, whereas the EMFO is employed for feature selection. Our suggested PRF is contrasted with recently applied COVID-19 recognition techniques in order to demonstrate its efficacy for identifying COVID-19 individuals such as (Maghdid et al., 2020; Sun et al., 2020; Nazish et al., 2021; Sen et al., 2021; Zhang et al., 2021; Habib et al., 2022; Khounraz et al., 2023). Results are shown in Tables 9 and 10. Tables 9 and 10 present the accuracy, precision, recall, F1-measure, macro-average precision, macro-average recall, micro-average precision, micro-average recall, and run-time for PRF to detect COVID-19 patients. The accuracy, precision, recall, F1-measure, macro-average precision, macro-average recall, micro-average precision, micro-average recall, and run-time are 98%, 93%, 92%, 76.9%, 92.49%, 91%, 88%, 89%, 86%, and 9 s, respectively, at training patient = 498. PRF also improves efficacy in terms of average, standard deviation, and median.

On the other hand, Fariba K. *et al.* (Khounraz et al., 2023) produce the worst performance because it is not sufficient with the limited number of images currently available about COVID-19 cases. Hence, the accuracy, precision, recall, F1-measure,

macro-average precision, macro-average recall, micro-average precision, micro-average recall, and run-time are 80%, 75%, 74%, 74.49%, 76%, 72%, 75%, 71%, and 26 s, respectively, at training patient = 498.

Although Sen S. *et al.* (Sen et al., 2021) have a higher performance than Fariba K. *et al.* (Khounraz et al., 2023), their performance is lower than that of (Maghdid et al., 2020; Sun et al., 2020; Nazish et al., 2021; Zhang et al., 2021; Habib et al., 2022), and PRF. The performance of Sen S. *et al.* (Sen et al., 2021) is reduced compared with other methods because it fails to disclose a considerable feature used for diagnosis. The technique in (Sen et al., 2021) is faster than in Sen S. *et al.* (Sen et al., 2021) and Fariba K. et al. (Khounraz et al., 2023) in training. Additionally, it is more flexible than Sen S. *et al.* (Sen et al., 2021) and Fariba K. *et al.* (Khounraz et al., 2023) in predicting COVID-19 cases. Although study in (Sun et al., 2020) has many advantages, it still has a longer training time than in (Maghdid et al., 2020; Nazish et al., 2021; Zhang et al., 2021; Habib et al., 2022), and PRF. The study in (Zhang et al., 2021) is worse than in (Maghdid et al., 2020; Nazish et al., 2021; Habib et al., 2022), and PRF because it requires a lot of memory and time.

Tables 9 and 10 show that the PRF's performance is significantly greater than in (Maghdid et al., 2020; Sun et al., 2020; Nazish et al., 2021; Sen et al., 2021; Zhang et al., 2021; Habib et al., 2022; Khounraz et al., 2023). Because FS^2 selects the most beneficial and trustworthy aspects for COVID-19 assessment, PRF bases its use on these features to swiftly and precisely identify COVID-19 cases that are affected. Lastly, according to several measures of assessment, PRF is far superior to other contemporary techniques in terms of detecting COVID-19 patients quickly and effectively. PRF is also more straightforward, adaptable, and capable of recognizing any condition. For identifying COVID-19 cases, the PRF has proven to be a trustworthy judgement system. It keeps the health care system from getting overburdened as a result.

The ANOVA test was used to examine the statistical difference between the proposed EMFO approach and the feature selection scenario. Another test, named the one-sample t-test, was also used for the

Table 15. The list of abbreviations used in this paper.

Abbreviation	Meaning
PRF	Patient's recognition framework
P^2S	Pre-processing stage
FS^2	Feature selection stage
CS	Classification stage
EMFO	Enhanced moth flame optimization
SVM	Support vector machine
WHO	World Health Organization
RT-PCR	Real-time polymerase chain reaction
GLCM	Gray Level Co-occurrence Matrix
MFO	Moth Flame Optimization
CNN	Convolution neural network
AUC	Area under curve
NLT	Numerical laboratory testing
MRI	Magnetic resonance imaging
CAP	Community-acquired pneumonia

Table 16. The list of symbols used in this paper.

Symbol	Meaning
$Moth_i$	i^{th} moth
$Flame_j$	j^{th} flame
S	Spiral function
D_i	Distance of the i^{th} moth for the j^{th} flame
b	The shape of the logarithmic spiral
t	A random number in [−1, 1]
l	Current number of iteration
N	Maximum number of flames
T	Maximum number of iterations
U	Non-empty set of finite objects
A	non-empty finite collection of features
$\gamma_R(D)$	Is the condition feature set R's classification accuracy in relation to the decision D
$\|R\|$	The length of the chosen feature portion.
$\|C\|$	The total number of features
β	Parameter corresponding to the subset length
α	Parameter corresponding to the importance of classification quality
P	Precision
R	Recall/sensitivity
A	Accuracy
E	Error

evaluation at a significance level of 0.05. For the COVID-19 detection scenario, the ANOVA test was used in the experiment to test the statistical difference of the proposed reference formulation (PRF). One more test, named the one-sample t-test, was also utilized in this part. This test can reveal whether the results of the proposed algorithm show a significant difference compared with those of other algorithms. The proposed PRF represents a statistically significant difference, according to the statistical analysis based on various tests. Finally, the list of abbreviations and the symbols used in this paper are illustrated in Tables 15 and 16, respectively.

7.1. Conclusion and future work

COVID-19 is one of the most serious issues that has brought life to a halt all over the world. Early recognition of COVID-19 patients is thus essential for the treatment and management of the disease. In this study, we offer a precise and efficient recognition system that might enable intelligent clinical assessment. In our recognition framework, PRF is composed of three stages: P^2S, FS^2, and CS. The P^2S extracts a set of features from CT scans for infected and healthy individuals. FS^2 chooses the most useful features from the features that have been obtained from P^2S by using the EMFO approach as a wrapper method. The CS employs the SVM classifier to efficiently identify suspected cases using the relevant features chosen by FS^2 within the shortest period of time.

According to analytical results, the suggested feature selection method outperforms other recent methods in terms of evaluation parameters while also being quick and effective. The EMFO approach has precision, recall, accuracy, and error values that are, accordingly, 0.92, 0.91, 0.94, and 0.06. The accuracy of the proposed PRF was 0.98, which is greater than that of other contemporary approaches. In terms of accuracy, precision, sensitivity, and execution time, the suggested PRF based on the EMFO approach and SVM delivers quicker and more accurate results than the recent techniques. The proposed framework has some limitations that need to be addressed. First, only the COVID-19 vs. normal classification challenge is used to validate the proposed PRF. Second, the suggested PRF relies mainly on a chest CT scan to evaluate COVID-19 detection. Third, the proposed PRF is limited to the detection of COVID-19. In the future, we want to apply our proposed framework to more COVID-19 classification tasks [e.g., COVID-19 vs. community-acquired pneumonia (CAP) and severe patients vs. non-severe patients]. In the future, we plan to apply the proposed PRF to different medical images for COVID-19, such as radiography and magnetic resonance imaging (MRI) scans. We also intend to address future illness-prognostic challenges, such as diverse kinds of cancer, Alzheimer's disease, and heart disease, using the proposed PRF.

Funding statement

No funds, grants, or other support was received.

Conflicts of interest

Declaration of conflicting interests statement: The author has no conflicts of interest to declare that are relevant to the content of this article.

References

Akl, A.A., Hosny, K.M., Fouda, M.M., Salah, A., 2023. A hybrid CNN and ensemble model for COVID-19 lung infection detection on chest CT scans. PLoS One 18, 3.

Alenezi, M., Alqenaei, Z., 2021. Machine learning in detecting COVID-19 misinformation on Twitter. Fut. Internet 13 (10), 244. https://doi.org/10.3390/fi13100244.

Ali, S.H., 2021. A new intrusion detection strategy based on combined feature selection methodology and machine learning technique. MEJ Mansoura Eng. Journal 46, 27–35.

Ali, S.H., El-Atier, R.A., Abo-Al-Ez, K.M., et al., 2020. A gen-fuzzy based strategy (GFBS) for web service classification. Wireless Pers. Commun. 113, 1917–1953.

Chengcheng, C., Qian, Z., Mahsa, H., et al., 2022. Forecast of rainfall distribution based on fixed sliding window long short-term memory. Engineering Applications of Computational Fluid Mechanics 16, 1.

Chiranji, L.C., Debi, P.A., 2020. Segmentation and feature extraction in medical imaging: a systematic review. Procedia Comput. Sci 176, 26–36. https://doi.org/10.1016/j.procs.2020.03.179.

Chung, M., Bernheim, A., Mei, X., et al., 2019. CT imaging features of 2019 novel coronavirus (2019-nCoV). Radiology 275, 202–207.

Deng, T., Huang, Y., Yang, G., Wang, C., 2022. Pointwise mutual information sparsely embedded feature selection. Int. J. Approx. Reason. 151 (c), 251–270.

Ghose, P., Alavi, M., Tabassum, M., et al., 2022. Detecting COVID-19 infection status from chest X-ray and CT scan via single transfer learning-driven approach. Front. Genet. 13.

Habib, S., Alyahya, S., Ahmed, A., et al., 2022. X-ray image-based covid-19 patient detection using machine learning-based techniques. Comput. Syst. Sci. Eng. 43, 671–682.

Huang, C., Wang, Y., Li, X., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395, 497–506.

Iancu, R.I., Zara, A.D., Mires tean, C.C., Iancu, D.P.T., 2022. Radiomics in COVID-19: the time for (R)evolution has came. BioMed 2, 60–68.

Jangir, N., Pandya, M.H., Trivedi, I.N., Bhesdadiya, R., Jangir, P., Kumar, A., 2016. Moth-flame optimization algorithm for solving real challenging constrained engineering optimization problems. In: 2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1–5. https://doi.org/10.1109/SCEECS.2016.7509293.

Kabir, M.M., Shahjahan, M., Murase, K., 2011. A new local search based hybrid genetic algorithm for feature selection. Neurocomputing 74, 2914–2928.
Kanagaraj, G., Kumar, P., 2020. Pulmonary tumor detection by virtue of GLCM. J. Sci. Ind. Res. 79, 132–134.
Kennedy, J., Eberhart, R.C., 1997. A discrete binary version of the particle swarm algorithm. In: International conference on systems, man, and cybernetics, computational cybernetics and simulation, 5, pp. 4104–4108. https://doi.org/10.1109/ICSMC.1997.637339.
Khounraz, F., Khodadoost, M., Gholamzadeh, S., et al., 2023. Prognosis of COVID-19 patients using lab tests: a data mining approach. Health Sci. Rep. 6, 1.
Krishanthi, G., Jayetileke, H., Wu, J., Liu, C., Wang, Y.-G., 2023. Enhancing feature selection optimization for COVID-19 microarray data. COVID 3, 1336–1355.
Kuzmenko, O., Lyeonov, S., Letunovska, N., Kashcha, M., Strielkowski, W., 2023. Impact of COVID-19 on the national development of countries: implications for the public health. PLoS One 18, 3.
Li, X., Yi, P., Wei, W., Jiang, Y., Tian, L., 2021. LNNLS-KH: a feature selection method for network intrusion detection. Secur. Commun. Netw. 3, 1–22.
Maghdid, H., Asaad, A., Ghafoor, K., Sadiq, A., 2020. Diagnosing COVID-19 Pneumonia from X-Ray and CT Images Using Deep Learning and Transfer Learning Algorithms. https://doi.org/10.48550/arXiv.2004.00038 arXiv:2004.00038.
Mahanty, C., Kumar, R., Mishra, B.K., Barna, C., 2022. COVID-19 detection with X-ray images by using transfer learning. J. Intell. Fuzzy Syst. 43, 1717–1726.
Mirjalili, S., 2015. Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm. Knowl. Base Syst. 89, 228–249.
Mirjalili, S., Mirjalili, S.M., Lewis, A., 2014. Grey wolf optimizer. Adv. Eng. Soft. 69, 46–61.
Mohan, G., Subashini, M., 2018. MRI Based medical image analysis: survey on brain tumor grade classification. Biomed. Signal Proc. Contr. 39, 139–161.
Naheed, N., Shaheen, M., Khan, S.A., Alawairdhi, M., Khan, M.A., 2020. Importance of features selection, attributes selection, challenges and future directions for medical imaging data: a review. Comput. Model. Eng. Sci. 125, 1.
Narin, A., 2020. Detection of covid-19 patients with convolutional neural network based features on multi-class X-ray chest images. In: 2020 Medical Technologies Congress (TIPTEKNO), pp. 19–20.
Nazish, Ullah, S.I., Salam, A., Ullah, W., Imad, M., 2021. COVID-19 lung image classification based on logistic regression and support vector machine. In: Musleh Al-Sartawi, A.M., Razzaque, A., Kamal, M.M. (Eds.), Artificial Intelligence Systems and the Internet of Things in the Digital Era. EAMMIS 2021. Lecture Notes in Networks and Systems, vol. 239. Springer, Cham.
Qinghua, Z., Qin, X., Guoyin, W., 2016. A survey on rough set theory and its applications. CAAI Transac. Intell. Technol. 1, 323–333.
Rostami, M., Berahmand, K., Forouzandeh, S., 2021. A novel community detection based genetic algorithm for feature selection. J. Big Data 8, 2.
Ruoyuan, Z., Yang, S., Xinghang, W., 2022. Network Intrusion Detection Scheme Based on IPSO-SVM Algorithm. 2022IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC). https://doi.org/10.1109/IPEC54454.2022.9777568.
Sadeghian, Z., Akbari, E., Nematzadeh, H., 2021. A hybrid feature selection method based on information theory and binary butterfly optimization algorithm. Eng. Appl. Artif. Intell. 97, 132–134.
Sen, S., Saha, S., Chatterjee, S., et al., 2021. A bi-stage feature selection approach for COVID-19 prediction using chest CT images. Appl. Intell. 51, 8985–9000.
Shanth, S., Rajkumar, N., 2021. Lung cancer prediction using stochastic diffusion search (SDS) based feature selection and machine learning methods. Neural Proc. Lett. 53, 2617–2630.
Sheela, M.S., Arun, C.A., 2022. Hybrid PSO–SVM algorithm for Covid-19 screening and quantification. Int J. Inf. Technol. 14, 2049–2056.
Sun, L., Mo, Z., Yan, F., Xia, L., 2020. Adaptive feature selection guided deep forest for COVID-19 classification with chest CT. IEEE J. Biomed. Health Inform. 24, 2798–2805.
Swathy, M., Saruladha, K., 2022. A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. ICT Exp. 8, 109–116.
Torse, D.A., Khanai, R., Pai, K., et al., 2023. Optimal feature selection for COVID-19 detection with CT images enabled by metaheuristic optimization and artificial intelligence. Multimed. Tool. Appl. 82, 41073–41103.
Vapnik, V.N., 1999. An overview of statistical learning theory. IEEE Trans. Neural Netw. 10, 988–999.
Vommi, A.M., Battula, T.K., 2023. A hybrid filter-wrapper feature selection using Fuzzy KNN based on Bonferroni mean for medical datasets classification. Expert Syst. Appl. 218. https://doi.org/10.1371/journal.pone.0282608.
Wah, Y.B., Ibrahim, N., Hamid, H.A., Abdul-Rahman, S., Fong, S., 2018. Feature selection methods: case of filter and wrapper approaches for maximising classification accuracy. Pertan J. Sci. Technol. 26, 1.
Wang, X., Yang, J., Teng, X., Xia, W., Jensen, R., 2007. Feature selection based on rough sets and particle swarm optimization. Pattern Recogn. Lett. 28, 459–471.
Xu, Y., Chen, H., Heidari, A.A., et al., 2019. An efficient chaotic mutative moth-flame-inspired optimizer for global optimization tasks. Expert Syst. Appl. 129, 135–155.
Zaim, S., Chong, J., Sankaranarayanan, V., Harky, A., 2020. COVID-19 and multiorgan response. Curr. Probl. Cardiol. 45, 1–22.
Zhang, Y.D., Satapathy, S.C., Liu, S., et al., 2021. A five-layer deep convolutional neural network with stochastic pooling for chest CT-based COVID-19 diagnosis. Mach. Vis. Appl. 32, 14.
Zhao, J., Zhang, Y., He, X., Xie, P., 2020. COVID-CT-Dataset: A CT Scan Dataset about COVID-19. https://doi.org/10.48550/arXiv.2003.13865 arXiv:2003.13865v1.
Zhu, R., Yu, Z.Y., Han, L., 2022. Insights on the possibility of SARS-CoV-2 transmission through the eyes. Int. J. Ophthalmol. 15, 1857–1863.
Zotin, A., Hamad, Y., Simonov, K., Kurako, M., 2019. Lung boundary detection for chest X-ray images classification based on GLCM and probabilistic neural networks. In: Proceedings of the 23rd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, 159, pp. 1439–1448, 314.