Informatics Bulletin, Faculty of Computers and Artificial Intelligence, Helwan University
Published Online ***Vol 5 Issue 2, July 2023***
(https://fcihib.journals.ekb.eg)

The Swarm Intelligence Algorithms for Optimization in Diagnosis the Diseases: A survey

Manar Elhussiny Ali[1], Osama E. Emam[2], Safaa M. Azzam[2]
[1]*Faculty of Computes and Information, Kafr El-Sheikh University, Cairo, Egypt.*
[2]*Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt.*
manarelhussiny2016@gmail.com , emam_o_e@yahoo.com , eng_safaa_azam@yahoo.com

***Abstract*— For quick diagnosis, treatment, and healing, many diseases need to be caught early. Delays in diagnosis can lead to other risks. Recently, researchers have been using artificial intelligence to find many diseases quickly and accurately. In particular, they have been using machine learning, CNN, and optimisation algorithms to pick the right features for a simple training model for the classification stage. As most data sets have noisy and repetitive features in all application areas, this slows down the performance of the classifier and may even make the classification less accurate because the search space is so big. This also affects the runtime of the classification. This review gives a full look at Particle Swarm Optimisation (PSO), Artificial Bee Colony (ABC), and Grey Wolf Optimisation (GWO). It will also talk about how these methods can be used to diagnose diseases like skin cancer, adrenal gland tumours, diabetes, coronary heart disease, and others. Also, the different procedures that researchers have taken to improve the accuracy and speed of diagnosis, the changes they have made to these algorithms, hybrids of these algorithms, and proposed future trends in every search The base of this study is to help new researchers get an overview of swarm intelligence algorithms and their role in diagnosing diseases and to brighten their horizons for the future directions in this field.**

***Keywords*— Feature selection, Metaheuristic algorithms, Swarm intelligence algorithm, Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), Grey Wolf Optimization (GWO).**

I. Introduction

The healthcare area is an application region for data mining since it has huge data resources that are hard to handle manually [1]. Medical data and patient records are used to uncover hidden patterns in disease diagnosis. Data mining and machine learning are used to retrieve the hidden knowledge within the data [2]. The prediction of an outcome based on historical data is one widely used machine learning application [1].

Machine learning is a subdomain of artificial intelligence [3]. One of the most important applications of machine learning is in the creation of disease regression models and categorization. [2]. Also, machine learning is used in many regions, including finance, retail, healthcare, and social data [4]. Machine learning started in the second half of the twentieth century. Samuel created the first machine-learning software in the early 1950s to play abstract strategy games like chess [5]. At the point when the forecast objective is continuous, regression is the fitting technique to utilise [3], where the instrument identifies patterns from the current dataset and then applies them to an unknown dataset to predict the outcome [1].

Classification is used when the goal is to get a single value or a class mark [3]. It is a reliable machine-learning process that is frequently used in forecasting. Some classification algorithms predict with pleasing accuracy, whilst others offer a restricted accuracy [1].

For classification, it is important to make a simple training model that is free of complexity and unimportant data. This brings us to the point of choosing features, which is to choose related attributes [6] that reduce the number of dimensions by getting rid of redundant or noisy features to choose a small subset of relevant features that work better, are more accurate, and cost less to compute [7]. Suppose dataset D contains N features where dataset D equals $A_1, A_2,.., A_N$. The starting point is to select the superior feature subsets from D. Evoke Subset d equals $A_1, A_2,.., A_n$ and, n < N and $A_1, A_2,.., A_n$ is the attributes of any dataset [6].

The metaheuristic algorithms were utilised to take care of the feature selection issue. Binary matrices Offers are evaluated for key characteristics. In the planned calculation, the arrangement vector is addressed. By [100101...], this means that 1 implies that an extraordinary attribute has been chosen, and a worth of 0 implies that the attribute isn't chosen in the subgroup [6].

The first section of this study will be an introduction. The second section will give some background information. The third section will talk about swarm intelligence algorithms. The fourth section will talk about how swarm intelligence algorithms can be used to diagnose diseases. The last section will be a conclusion.

II. Background

Feature selection (FS) is one of the most important problems in pattern recognition. Its goal is to reduce the number of features used for recognition while still getting a good result [8]. The FS algorithm picks out a subset of the original set of features that are important and enough for classifying. The classification process will be easier, faster, and more accurate as a result of this reduction. In addition, because of this reduction, the identification of features that do not need to be stored, collected, or bought may bring financial savings [8]. The classification of feature selection methods is shown in figure 1 [6].

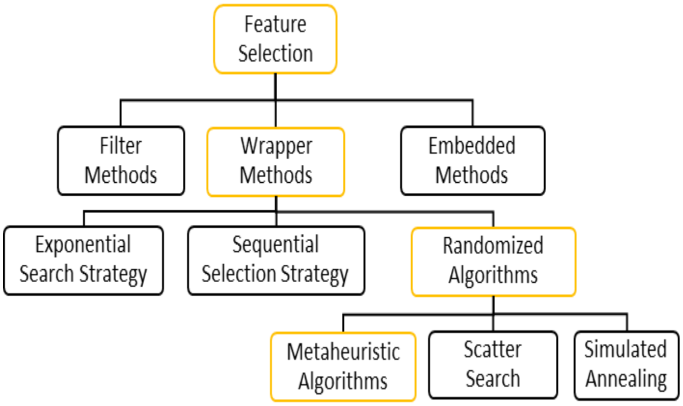

Fig. 1 Classification of feature selection methods

This research will deal with one type of feature selection, which is metaheuristic algorithms, especially swarm intelligence algorithms. The word "metaheuristic" is made up of the words "heuristic" and "meta." Since "heuristic" means to return to or find a goal through trial and error, and "meta" means "higher level," "metaheuristics" means "a higher level of heuristics" [9]. Various historical periods have progressed meta-heuristics or heuristic algorithms. Heuristic approaches were widely employed in many applications during the 1940s and 1960s, but the development of evolutionary algorithms in 1963 was an evolutionary algorithm (EA). View the development of genetic algorithms in the 1960s and 1970s. Simulated annealing was one of the most important milestones in metaheuristic algorithms in the 1980s and 1990s. The invention of various major algorithms such as ACO, PSO, and differential evolution during the 1990s and 2000s was an exciting time for metaheuristic algorithms. Metaheuristic algorithms have been widely employed in a variety of applications since 2000, and several new noteworthy algorithms have been created [9]. Metaheuristics are a collection of solutions that are too large to sample completely. Metaheuristics may generate a few hypotheses regarding the optimisation problem being solved, making them suitable for a wide range of issues. When it comes to optimisation methods, metaheuristics do not imply that the best or globally optimum solution may be discovered in a variety of situations. Many metaheuristics use stochastic optimisation to find a solution that is dependent on the set of random variables produced [9].
The attributes of most metaheuristics [9] are:

1. Metaheuristics are search-guiding algorithms.
2. The goal is to survey the search area in its proper location to find solutions that are close to ideal.
3. Techniques for forming metaheuristic algorithms, ranging from simple local search stages to complex learning procedures.
4. Metaheuristic algorithms are non-deterministic and approximate in nature.
5. Metaheuristics are a separate issue.

The two basic groups of metaheuristic algorithms are as follows [6].

1. *Single-solution metaheuristic algorithms:*

These mechanisms start with a single solution and modify it as they move through the optimisation process. It may cause local optima to be trapped, and it also fails to recognise the whole search area [6].

2. *Multiple-solution metaheuristic algorithms:*

These algorithms start by creating a population of solutions and then optimising them. With more repetitions, the population of solutions changes. Multiple solutions aid each other and have a massive exploration of search space; therefore, the methods are useful for avoiding local optima. They also possess the ability to jump to the most promising area of the search space. Therefore, population-based algorithms are used in solving most real-world problems [6].

The other categorization for stochastic optimisation techniques is based on their motives and inspirations. which are divided into four groups [10], as illustrated in figure 2 [6].

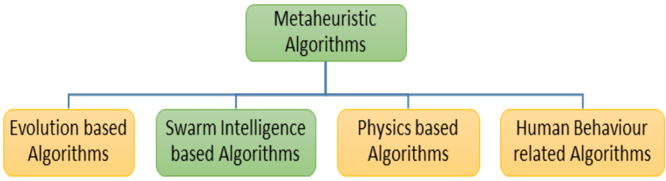

Fig. 2 Classification of metaheuristic algorithms

1. *Evolutionary-based (EB)*

It's a population-based metaheuristic that mimics biological evolution processes including selection, mutation, crossover, recombination, and chemotaxis. Genetic algorithms (GA), scrambled frog leaping algorithms (SFLA), and differential evolution (DE) are some of the most well-known EBs [4].

2. *Physics-based (PB)*

Physical laws in nature, such as gravitational force, electromagnetic force, and energy conservation, are generally used to manage and improve a population. The general mechanism of PB is different from that of other approaches because agents communicate and share information based on physical rules. In PB, SA and the gravitational search algorithm (GSA) are two representations. The big bang-big crunch algorithm (BB-BC), wind-driven optimisation (WDO), central force optimisation (CFO), and galaxy-based search algorithm are some of the other upcoming PBS [4].

3. *Human-based (HB)algorithms*

They're a freshly discovered category in intelligence computing that mathematically stimulates human social actions to generate near-optimal solutions. Some of the HBs are social group optimisation (SGO), cultural evolution algorithms (CEA), and ideology algorithms (IA) [4].

4. Swarm-based (SB)

It always encourages social beings to cooperate to find better solutions to challenges. Natural colonies, flocks, and herds are often the source of inspiration for SBs. Particle swarm optimisation (PSO), ant colony optimisation (ACO), and artificial physics optimisation (APO) are three of the most well-known SBs [4].

III. SWARM INTELLIGENCE ALGORITHMS

"Any attempt to build algorithms or distributed problem-solving devices based on how social insect colonies and other animal societies work as a group," says Bonabeau. The phrase "swarm," on the other hand, is commonly used to describe any small collection of interacting persons or agents. Bees swarming around their hive is a classic illustration of a swarm [11]. It is critical to determine whether a behaviour is swarming smart behaviour or not in order to construct a revolutionary swarm intelligence-based algorithm. suggested that the division of labour and self-organization are necessary and sufficient prerequisites for achieving intelligent swarming behaviours [12].

A. Self-organization

It consists of a set of dynamic mechanisms. These processes establish the primary rules for system component interactions [11]. This is an important aspect of a swarm frame because it results in a global standard constraint through the ability to interact among its low-standard components without the need for a central authority or external element to enforce it through planning. As a result of the local interaction of the components that make up the frame, the globally unified style emerges, and the organisation is done in parallel since all of the items act at the same time and are distributed because no item is a central coordinator [12]. There are four key characteristics of self-organization:

- Positive feedback: When data from a system's output is put back into one of its inputs, it encourages the building of structures that are useful. In the field of swarm intelligence, positive feedback makes the system more interesting and moves it to a new stable state [12].
- Negative feedback: serves to sustain the collective pattern by compensating for the influence of positive feedback [12].
- Fluctuations: the pace or amplitude of random system changes. For efflorescent structures, randomness is important because it allows for the finding of new solutions. It aids in the foraging process by removing stagnation [12].
- Multiple interactions: allow individuals within a community to learn from one another, boosting collective swarm intelligence [12].

B. Division of labour

It entails teamwork on specified activities and in comparable roles. Different tasks are carried out concurrently by specialised individuals in a group. Collaborative expert persons' simultaneous task performance is thought to be more active than unexpert individuals' sequential task performance [12].

C. A smart swarm must have five principles:

- the proximity: The swarm has to be eligible to do simple time and space computations [13].
- the quality: The swarm must be able to react to quality parameters in the environment [13].
- diverse response: the swarm doesn't have to confine its actions to confined pathways [13].
- stability: The swarm does not have to change its behaviour in response to environmental changes [13].
- Adaptability: The swarm must be eligible to change behaviour mode as needed [13].

Swarm intelligence (SI) is the mass behaviour of self-organized and decentralised swarms, such as fish groups, bird flocks, and social insect colonies such as bees, ants, and termites [13]. We will discuss some of these algorithms in detail below.

1. Particle Swarm Optimization (PSO)

Dr. R.C. Eberhart, an electrical engineer, and Dr. James Kennedy, a social consultant, developed a random optimisation approach that became known as particle swarm optimisation in 1995 while seeking to recreate the well-choreographed, beautiful movements of a flock of birds as part of socio-cognitive research on collective intelligence in biological communities [14]. It is a heuristic for universal optimisation. It is at present one of the most commonly used optimisation techniques [15]. Although PSO can resolve intricate, multi-distinct problems a lot more effectively than classic methods and numerous new optimisation methods, its efficiencies in solving the problems are still not as good as foreseeable [16]. It is a fashion for optimising robust numerical functions using the metaphor of the social attitude of flocks of birds and groups of fish. A swarm consists of individuals, called particles, and every particle represents a potential solution to the problem [16]. PSO uses a velocity vector to update the current position of every particle in the swarm. Each particle's position is updated based on the social behaviour of a population of people, the swarm in the case of PSO, which adjusts to its surroundings by returning to previously found promising zones. The processes are stochastic in nature and rely on each particle's memory as well as the collective knowledge of the swarm [17].

A. A diagram of a simple PSO algorithm

- begins with a swarm initialization, which is generally spread randomly over the search area [17].
- For each member in the swarm, calculate a velocity vector [17].
- Using the prior location and the updated velocity vector, update the position of each person [17].
- Repeat step 2 until convergence is achieved [17].

B. The mathematical model for SPO

The equation for altering the position of every individual is [17]:

$$\mathbf{X}_{k+1}^{i} = \mathbf{X}_{k}^{i} + \mathbf{v}_{k+1}^{i}\, \Delta \mathbf{t} \qquad (1)$$

Individual i's location at repetition k+1 is X_{k+1}^{i} and v_{k+1}^{i} is the associated velocity vector. Throughout this paper, a unit time step t is employed [17].
The equation for updating each individual's velocity vector is determined by the PSO algorithm in question [17]. The following is a frequent equation:

$$\mathbf{v}_{k+1}^{i} = \mathbf{w}\mathbf{v}_{k}^{i} + \mathbf{c}_1\mathbf{r}_1\frac{(\mathbf{p}^{i} - \mathbf{x}_{k}^{i})}{\Delta \mathbf{t}} + \mathbf{c}_2\mathbf{r}_2\frac{(\mathbf{p}_{k}^{g} - \mathbf{x}_{k}^{i})}{\Delta \mathbf{t}} \qquad (2)$$

Where r_1 and r_2 are independent random numbers between 0 and 1, individual i has discovered the best position p^{i} so far, and the best position in the swarm at time k is p_{k}^{g}. A time interval step Δt is employed once again is used throughout the current work. The individual's inertia w and two "trust" factors c_1 and c_2 are problem-dependent parameters [17]. We see a chart of the PSO algorithm in Figure 3 [18]:

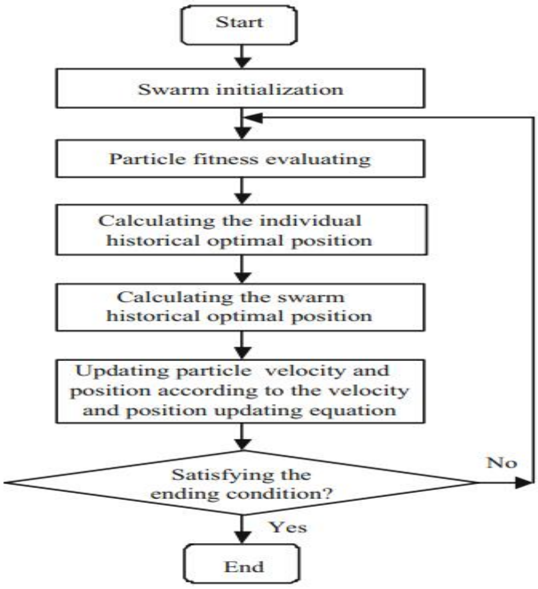

Fig. 3 flowchart of the particle swarm optimization algorithm.

PSO is defined by a number of variables that play a crucial part in determining the optimum solution, including:

1. Inertia Weight

In the search space, the speed of an individual in each dimension is fixed at the maximum speed, v_max. This maximum intent speed determines which areas between the current position and the target position can be sorted. In experimental work, w is kept between 0.9 to 0.4, and the values are reduced linearly to reach the optimal target position quickly [14].

2. Acceleration constants c1 and c2

It regulates the length and time it takes for the particles to reach the optimum position. These constants must be chosen correctly. In general, each of these constants is set to 2 to make the time taken to travel toward the particle's subjective best and global swarm best equal to half the total time [14].

3. Random numbers r1 and r2

Random numbers are added to the update algorithms to control particle drag towards the pbest and gbest places. They provide the PSO algorithm a random component, which helps it avoid becoming trapped at a non-ideal local minimum or maximum [14].

4. Size of the population

This is frequently classified experimentally based on the dimensions and perceived difficulty of a task. Values in the 20–50 range are rather typical. The size of the swarm changes from application to application, and is hence issue-dependent [14].

C. Advantages of PSO

- PSO The PSO algorithm does not account for selection or mutation computations. The search may be carried out by varying the particle's speed repeatedly [14].
- Particles only fly to excellent regions (where there is a chance of obtaining food) after learning from the group's experiences [14].
- Because the PSO method is based on artificial intelligence, it may be used in both scientific and engineering applications [14].
- The PSO method requires simple computations, which are becoming easier to perform with the introduction of better evaluation techniques [14].

D. Disadvantages of PSO

- With increasing search space dimensions, standard PSO suffers from a significant increase in search complexity [14].
- The approach is susceptible to partial optimism, which results in substantially less precise speed and direction regulation [14].
- This approach cannot be used for non-coordinate system issues, such as the solution to the energy field and the movement rules of the particles in the energy field, due to a lack of dimensionality [14].

2. Artificial Bee Colony (ABC)

It is a branch of swarm intelligence [19]. that is derived from honey bee behaviour in swarms consists of The model of forage selection of mass intelligence of bee swarms which consists of three main components: unemployed foragers, employed foragers, and food source [11].

- Food Sources: The value of a food supply is determined by a variety of parameters, including its proximity to the mite, the density of its energy, and the ease with which it may be extracted [11].

- Employed foragers: They're related to a certain food supply that they're taking advantage of. They save information on this particular source, its location and direction from the mite, as well as the source's profitability, and they communicate this information with a specific probability [11].
- Unemployed foragers: are responsible for constantly looking for a food source. There are two unemployed foragers: onlookers biding their time in the nest and establishing an eating source through the information given by employed foragers; and scouts screening the environment around the nest for novel eating sources [11].

This model distinguishes between two main forms of behaviour: mobilisation to a nectar source and desertion from the source [11]. The most crucial manifestation in the construction of mass knowledge is the exchange of information among bees. The dancing zone is the most significant part of the hive in terms of information exchange. The dancing area is where bees communicate about the quality of their food sources. As depicted in figure 4, this dance is known as the waggle dance [11].

Fig. 4 waggle dance of bees.

ABC's search strategy is inspired by bee foraging activities. In the solution space, food sources that are scattered across the environment are referred to as solutions. Scout bees, spectator bees, and employed bees are the three types of bees in a swarm. The employed bees sift through the available options in an attempt to choose a few that are the best. The search information is shared across all bees. The observer chooses the best solutions and does a further search. The scout bees keep an eye on the swarm's various alternatives. In certain repetitions, the scout bees create a random solution to replace the unchanged answer [19].

A. The mathematical model for ABC

The population of food sources is established by artificial scout bees during the startup phase, and control settings are set [13]. The initial swarm has SN solutions, and the swarm size is SN. The following is how each initial solution X_i is generated at random [19]:

$$\mathbf{x_{i,j} = low_j + rand \cdot (up_j - low_j), i = 1, 2, \cdots, SN, j = 1, 2, \cdots, D,} \tag{3}$$

D is the dimension size, while rand is a random number between 0 and 1, and $[low_j, up_j]$ is the boundary constraint [19]. During the employed bees phase, artificially hired bees hunt for novel food sources with more nectar associated with the food source in their memory. They find a local food source and determine its adequacy. After the new food source is created, its fitness is judged, and a grasping choice is made between it and its parent. The hired bees then dance in the dancing area to relay their eating source information to the hive's spectator bees [13]. The hired bees scour each solution X_i for a new one V_i [19]:

$$\mathbf{v_{i,jr} = x_{i,jr} + \emptyset_{i,jr} \cdot (x_{i,jr} - x_{k,jr}), i = 1, 2, \cdots, SN,} \tag{4}$$

where X_k is a random number drawn from the swarm (($X_k \neq X_i$), and jr ∈ [1, D] is a random integer. The weight $\emptyset_{i,jr}$ is chosen at random between -1 and 1. As can be observed, Eq. (2) only affects X_i 's jrth dimension. Both V_i and X_i have the same values for the remaining dimensions. X_i is replaced by V_i when V_i is superior than its parent X_i [19].

During the onlooker bee phase, onlooker bees select their food sources based on information provided by employed bees. Use a fitness-based selection strategy for this goal. After deciding on a feeding source for an observer, the bee is randomly selected, a nearby source is found, and its fitness value is determined 2 [13].

Onlooker bees' work differs from that of hired bees. They don't look for all answers, but rather the finest ones. Each X_i has a selection probability P_i according to the fitness proportional selection (FPS) approach. In terms of P_i, many best answers are chosen, and the observer bees do an additional search on these. [19] computes the probability P_i :

$$\mathbf{P_i = \frac{fit(X_i)}{\sum_{i=1}^{SN} fit(X_i)},} \tag{5}$$

$$\boldsymbol{fit(X_i) = \begin{cases} \frac{1}{1+f(x_i)}, & if\ f(x_i) \geq 0 \\ 1 + |f(x_i)|, & otherwise \end{cases}} \tag{6}$$

The fitness and function values of X_i are $fit(X_i)$ and $f(x_i)$, respectively. When the selection probability P_i for each solution X_i in the present swarm is met, the spectator bees employ Eq. (4) to generate offspring V_i. Onlooker bees, like the hired bees, utilise greedy selection to compare the quality of V_i and X_i. The best of the two is selected as the new X_i. Scout bees use a counter to keep track of the changes in each solution in the swarm throughout their search phase. If X_i cannot be substituted by V_i, the counter $trial_i$ is increased by one; otherwise, $trial_i$ is set to zero. Eq. (3) [19] re-initializes the associated solution X_i when $trial_i$ exceeds a predetermined value limit. Figure 5 [13] shows the ABC optimization approach's general algorithmic structure:

```
Initialization Phase
REPEAT
Employed Bees Phase
Onlooker Bees Phase
Scout Bees Phase
Memorize the best solution achieved so far
UNTIL (Cycle = Maximum Cycle Number or a Maximum CPU time)
```

Fig. 5 The general algorithmic structure of the ABC optimization

B. Studies on ABC optimization

This research [13] provides a summary of studies on ABC. The revisions to ABC and publications linked to ABC in terms of application areas are provided in the latter [13]. Also studied in this study [11] is a special intelligent method of a bee swarm's foraging manner and a revolutionary artificial bee colony algorithm that imitates this foraging manner of actual bees for solving multimodal and multidimensional optimization problems [11].

C. Disadvantages of ABC

The disadvantages of ABC are that it does not use an operator-like crossover like GA or DE, and the distribution of excellent information among solutions is not as good as it should be. The ABC convergence performance for the local minimum suffers as a result of this. A highly significant practical problem is symbolic regression, which is the process of obtaining a mathematical model by a defined limit sampling of values of distinct variables and connected values of dependent variables. So far, ABC has not been used to solve the symbolic regression problem [13].

3. Grey Wolf Optimization (GWO)

The Canidae family includes the grey wolf. Grey wolves prefer to live in packs. The average group size is 5–12. They have a strict social control hierarchy in place [20]. Alphas, whether female or male, are the leaders. The alpha is in charge of deciding on things like wake time, hunting, sleeping location, and so on. The bundle is ordered by the alpha's judgments. Several democratic attitudes have been recorded, in which the alpha wolf follows the other wolves in the pack. Only the alpha wolves are permitted to mate outside of the bundle. The betas are subordinate wolves who assist the alpha in making decisions and other important activities. The beta wolf can be either female or male, and it is a better contender to be the alpha wolf if one of the alpha wolves dies or grows old. It also gives commands to those who are lesser wolves. A scapegoat is the omega. All other controlling wolves must always be subordinate to the Omega wolves. They are the last wolves who will devour anything. It may appear that the omega is an insignificant member of the bundle, yet it has been noted that the entire bundle confronts indoor conflict and issues as a result of the omega's loss. Beta and alpha wolves should be subjected to delta wolves, but they have influence over the omega. Scouts, sentinels, elders, and others fall within this group [20].

A. the major grey wolf hunting phases are as follows [20]:

- Tracking, chasing, and, oncoming the prey.
- Chasing, surrounding, and discomfort the prey to it stop moving.
- Attack the prey.

B. The mathematical model for GWO

It contains five phases the first phase is a social hierarchy. The best answer is designated as the alpha. The second and third-best solutions are thus referred to as beta and delta, respectively. The remaining possible solutions are all expected to be omega. The hunting (optimization) in the GWO algorithm is directed by, and. These three wolves are being pursued by the wolves. Grey wolves encircle prey during the hunt in the second phase of encircling prey. The following equations [20] are presented to quantitatively represent encircling behaviour:

$$\vec{\mathbf{d}} = |\vec{\mathbf{v}} \cdot \overrightarrow{\mathbf{X_p}}(\mathbf{i}) - \vec{\mathbf{X}}(\mathbf{i})| \tag{7}$$

$$\vec{\mathbf{X}}(\mathbf{i}+\mathbf{1}) = \overrightarrow{\mathbf{X_p}}(\mathbf{i}) - \vec{\mathbf{m}} \cdot \vec{\mathbf{d}} \tag{8}$$

where i means the current repetition, $\vec{v}$ and $\vec{m}$ are coefficient vectors, $\overrightarrow{X_p}$ is the situation vector of the prey, and $\vec{X}$ means the situation vector of a grey wolf. The vectors $\vec{m}$ and $\vec{v}$ are calculated as follows [20]:

$$\vec{\boldsymbol{m}} = 2\vec{\boldsymbol{l}} \cdot \overrightarrow{\boldsymbol{r_1}} - \vec{\boldsymbol{l}} \tag{9}$$

$$\vec{\boldsymbol{v}} = 2 \cdot \overrightarrow{\boldsymbol{r_2}} \tag{10}$$

where r_1, r_2 are random vectors in [0, 1], and components of an are linearly reduced from 2 to 0 along the track of repeats. to show how Eqs. (9 and 10) interact. A grey wolf in the circumstance (X, Y) can modify its location in response to the prey's situation (X*, Y*). Various locations around the optimal factor can be obtained by altering the values of $\vec{m}$ and $\vec{v}$ vectors while taking the present circumstance into account. For example, by putting $\vec{m}$=(1,0) and $\vec{v}$=(1,1), you may get (X* –X, Y*). In 3D space, the various updated circumstances of a grey wolf. Note that wolves can attain any scenario between the locations [20] using the random vectors $r_1$and r_2.

Hunting is the third phase. Grey wolves have the capacity to recognise and surround prey. The alpha is in general charge of the hunt. When hunting, the delta and beta could work together. However, they have no knowledge of the ideal site in a summary search space. They suppose that the alpha (the best solution), beta, and delta have greater learning about prospective prey sites in order to mathematically emulate the hunting approach of grey wolves. As a result, they preserve the first three best answers obtained so far and instruct the remaining search agents to adjust their locations in order to benefit from the better search agents' placements. [20] The following equations are suggested:

$$\overrightarrow{d_\alpha} = \left|\overrightarrow{v_1} \cdot \overrightarrow{X_\alpha} - \vec{X}\right|, \overrightarrow{d_\beta} = \left|\overrightarrow{v_2} \cdot \overrightarrow{X_\beta} - \vec{X}\right|, \overrightarrow{d_\delta} = \left|\overrightarrow{v_3} \cdot \overrightarrow{X_\delta} - \vec{X}\right| \tag{11}$$

$$\overrightarrow{X_1} = \overrightarrow{X_\alpha} - \overrightarrow{m_1} \cdot \left(\overrightarrow{d_\alpha}\right), \overrightarrow{X_2} = \overrightarrow{X_\beta} - \overrightarrow{m_2} \cdot \left(\overrightarrow{d_\beta}\right), \overrightarrow{X_3} = \overrightarrow{X_\delta} - \overrightarrow{m_3} \cdot \left(\overrightarrow{d_\delta}\right) \tag{12}$$

$$\vec{X}(i+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{13}$$

The fourth step involves catching prey (exploitation). Grey wolves finish the hunt by striking the victim when it stops moving, as previously stated. They diminish its value by mathematically modelling approaching the prey. It's worth noting that 's fluctuation range has shrunk by . In other words, is a random value in the range [-2l, 2l], where a decreases from 2 to 0 during the course of the repetitions. When has random values in [-1,1], a search agent's future circumstance might be anywhere between its current position and the prey's situation [20].

The search for prey is the fifth phase (exploration). Grey wolves search primarily s the alpha, beta, and delta positions. They move apart to examine for prey and then converge to assault it. They employ A with random values less than -1 or larger than 1 to direct the search agent to splay from the prey in order to mathematically describe divergence. This indicates that the GWO algorithm can search the whole world [20]. Figure 6 shows a pseudocode for the GWO algorithm.

```
Initialize the grey wolf population Xi (i = 1, 2, ..., n)
Initialize a, A, and C
Calculate the fitness of each search agent
Xα=the best search agent
Xβ=the second best search agent
Xδ=the third best search agent
while (t < Max number of iterations)
    for each search agent
            Update the position of the current search agent by equation (3.7)
    end for
    Update a, A, and C
    Calculate the fitness of all search agents
    Update Xα, Xβ, and Xδ
    t=t+1
end while
return Xα
```

Fig. 6. Pseudo code of the GWO algorithm.

C. Studies on GWO optimization

Researchers have made many studies and improvements to GWO in the different fields, and some of these improvements are mentioned in this review [21], while studies in the feature selection field are mentioned below.

Some researchers made different methods using GWO and KNN as classifiers, but their works relied on very tiny population sizes, ranging from 5 up to only 8, and they compared their approach with PSO and GA over different benchmark datasets. The first method combines a k-nearest neighbour (k-NN) fitness function with a GWO binary version to assess the specified subsets of features. Experiments were done on 18 datasets using the same wrapper design with a k-NN fitness evaluator as in their earlier work. and others developed a method that consists of two primary stages: the first is a filtered-base approach that uses mutual information equations as a fitness function, and the second is a wrapper-based technique that uses a classifier as an evaluator. Another approach was made for GWO binary to classify cancer on gene expression data, and they used a decision tree algorithm in this approach. In this method, they depended only on accuracy without concern for the number of features, and it was tested based on 10 cancer datasets and other classifiers such as SVM and MLP. A feature-reducing approach was proposed to GWO to search the feature space for a subset of features that raise the classification fitness function. They also took into account the accuracy rate. This approach was tested on 11 datasets and compared with conditional entropy-based attribute reduction (CEAR) and other methods [21]. For more studies on GWO, check out this review [21].

D. Disadvantages of GWO

When tackling real-world problems, GWO may need to be tweaked. because no single optimization technique can tackle all optimization issues. Because of its single-objective nature, this method can only solve single-objective problems. The fundamental disadvantage of GWO is its limited ability to deal with the challenges of a multi-modal search environment. The GWO algorithm's efficiency declines considerably as the number of variables increases [21].

IV. SWARM INTELLIGENCE ALGORITHMS IN DIAGNOSIS THE DIFFERENT DISEASES

In this section, we will discuss the uses of PSO, GWO, and ABC in diagnosing different diseases, the different procedures that researchers have taken to improve the accuracy and speed of diagnosis, the changes they have made to these algorithms, hybrids of these algorithms, and proposed future trends in each search.

A. Particle Swarm Optimization (PSO)

In this study [22], they came up with a way to spot early skin cancer. If the disease isn't treated early, it could spread to the area below the skin, making it hard to treat. In the first stage, feature extraction, 1000 features are taken from a data set of treated skin cancer images from the Kaggle website. This data set has 3297 images, 1497 of which are malignant and 1800 of which are benign. 1000 features were taken from this data set using efficientNetB0, and then we moved on to the second stage, feature selection, where the characteristics were cleaned up using PSO and GA to get the required features. In both ways, they used KNN to figure out the fitness value, and in the third step, they put all the results together. Put the results of both methods into one entry, and then use SVM to classify the entry. This model achieved an accuracy of 89.17%. In future work, they suggest the use of other CNN technologies and different feature selection algorithms [22].

Others used BPSO, formed from PSO, to explore the value of magnetic resonance imaging for diagnosing adrenal gland tumours. The data used in this method were collected from 120 patients with adrenal gland tumours, which contained 52 patients with malignant tumours and 68 others with benign tumours. They were further divided into four groups according to where the tumours were located: metastatic, stromal, medullary, and cortical. These patients were randomly divided into two groups, each with 60 patients: the control group, which was examined using conventional MRI, and the observation group, which was examined using MRI in conjunction with BPSO, with the second group achieving an accuracy of 81.67 percent and the first group achieving an accuracy of 58.33 percent. They pointed out that the results were not sufficient, and the relationship between the site of the disease and benign and malignant tumours was not studied as a future direction for research [23].

In this research [24], they made a model that classifies the types of diabetes accurately based on a dataset collected from the Palestinian Diabetes Institute, where the DataPal consists of nine features used to diagnose types of diabetes and 314 female cases, and also contains both types of diabetes (T1DM and T2DM), and this data was processed using KNN to supplement the data. For this method, PSO and multi-layer perceptron neural networks (ML-PNNs) were combined, and the PSO coefficients were also improved before the two networks were combined. This model was accurate 98.73 percent of the time and made suggestions for the future, such as making a medical app that helps people with a family history of diabetes by using machine learning algorithms to give them advice and treatments. Using the fuzzy rule with the PSO-MLPNNs model to predict diabetes and automatically give treatments, building a prediction and classification system with real-time remote monitoring applications [24].

Others proposed an effective method for predicting heart disease using a stacked sparse autoencoder with a softmax layer where the hidden layer of another stacked sparse autoencoder is connected to a softmax layer classifier that consists of an SSAE network. They used PSO to improve how well the SSAE feature learned and sorted information. The proposed method was tested on two datasets. The first is in Cleveland, and the second is in Framingham, where the first group contains 303 people, including 165 patients and 138 healthy people, and the other group contains 4238 people, including 3594 patients and the rest 644 healthy, but both datasets contain missing values, so they used the mean imputation to deal with the missing data, as this method is based on replacing missing data with a substitute value based on other available details. An accuracy of 97.3% was reached for the Cleveland data set and 96.1% for the Framingham data set. Their future trend was to study the effect of feature learning on different classification algorithms and to focus on stacking different variants from autoencoders [25].

B. Grey Wolf Optimization (GWO)

In recent years, the number of people with diabetes has gone up by a large percentage. This is because more people are living unhealthy lives and don't move around enough, among other things. To help predict diabetes because early detection reduces side effects, they made a model. Datasets related to diabetes were obtained from the UCI machine learning repository, and the dataset contains 8 features and 2 classes. This model creates 17 fuzzy rules using features, and then these rules are provided as input into the GWO algorithm, which provides the optimal rules for the output. The proposed work was compared with the ant colony optimisation algorithm, where the proposed work gave an accuracy of 81% and the ant colony optimisation algorithm gave an accuracy of 71% [26].

Also, early diagnosis of rheumatoid arthritis is a challenging task for general practitioners who have developed a method for the effective prediction of rheumatoid arthritis. Using PSO to select primitive positions and then GWO to update the current positions of the population from the search space to get the optimal features for better classification, the selected features were then presented as inputs for the C4.5 approach. The proposed model was examined based on real-time patient data, and the proposed model was compared with two newer approaches, including CSBOOST and REACT, and the proposed approach was 10 times better than the other approach. This model reached an average accuracy of 86.36 percent, and they suggested in a future study that it may include the inclusion of patient details in real-time and applying different approaches of ML to improve prediction performance [27].

Others improved the grey wolf's ability to recognise early-stage Parkinson's disease symptoms, and the proposed model was evaluated using three data sets. The first group was for HandPD, as it was prepared from 158 individuals, including 105 patients and the rest in good health, and it consisted of 632 cases and 13 features. The second group was for SpeechPD, as it was prepared from recordings of 31 individuals, including 23 patients. A total of 195 cases and 239 features were created. The third VoicePD group was prepared from the voices of 20 patients and 20 healthy individuals, containing 1040 cases and 26 features, and compared between MGWO and OCFA, and it has a higher accuracy of approximately 94.83%. The results showed that MGWO's use of random forests outperformed other classifiers in terms of accuracy. In order to further study this, we formulate new methods for integrating voice, speech, and hand data models so that early detection is carried out with better accuracy, and we apply the proposed method to it and use the proposed method to solve other improvement problems [28].

They used brain image analysis to discover Alzheimer's disease. Initially, the images are preprocessed to remove unwanted areas using a Gaussian filter, and after denoising all the data, the features are extracted from the images. I used a grey-level co-occurrence matrix (GLCM) and a grey-level run length matrix (GLRM) and other techniques at this stage. Use group grey wolf optimisation (GGWO) to choose the optimal features for this problem, and finally, the classification process is carried out using different classifiers such as KNN, decision trees, and neural networks (NN). They used a data set consisting of 287 cases and three classes: ordinary, mild cognitive impairment, and Alzheimer's disease (AD). To assess the proposed model's 96.23% accuracy [29].

C. Artificial Bee Colony (ABC)

Because Parkinson's disease is a neurological disease that cannot be fundamentally treated and because early automatic detection of it in data sets is a difficult task today, they proposed a model consisting of three stages. The first stage is the filter, which uses three algorithms to sort and weight the features. The outputs extracted from this stage are inputs to the second stage, called the evaluation score, which ranks these results with normalisation to obtain a comprehensive classification of features, and the third stage is called the wrapper. It uses the new ranking as the basis for the population. The first two MFABC algorithms assign a weight to each feature according to its rank, and the feature with the largest weight has the greatest probability of selection and is then used as an input to the classifier. This algorithm was investigated using five different datasets for Parkinson's disease. The accuracy of the PD speech 2 datasets, the HandPD spiral dataset, the PD acoustic dataset, the PD speech dataset, and the HandPD meander dataset are 94.05%, 97.62%, 100%, 100%, and 96.83%, respectively [30].

And also look at Malaria is a life-threatening disease in Africa, where the World Health Organization reported that an estimated fifty million children died of malaria within five years, so they created a model for diagnosing malaria based on symptoms, which is a web application consisting of three links: the home page, patient registration, and patient diagnosis request. They used the ABC and structured analysis and design technique (SADT). This technique is a structured analysis modelling language that uses two types of activity graphs and data model graphs that are specifically used to meet group requirements. They also implemented it using MySQL and PHP. They took different samples and predictions to evaluate the proposed model, and in general, 96% of correct predictions and 4% of false predictions were found. They compared the present and improved malaria diagnosis models, and the improved model proved efficient in terms of accuracy, treatment speed, classification error, the number of symptoms treated, and area under the curve (AUC) [31].

In addition, others created a model called ABC-FS-Optimized DNN to more accurately diagnose stroke patients. They used DNN to deal with the imbalance of known stroke data, as it contains 43,400 cases, including 42,617 healthy individuals and the rest suffer from stroke. The data set contains many missing values, and these missing values are replaced by average values for each feature. ABC was relied on to identify the relevant features of stroke disease and then take these features as input to the CNN model. This technique consists of three layers; the first layer is responsible for scanning the input data using the kernel, and each kernel represents a feature from the dataset. The next layer is responsible for the activation process of neurons based on the information, and if the information is suitable, the activation is done; otherwise, the activation is deactivated, and the last layer combines the extracted features to create a feature set. This model achieved 87.09% accuracy [32].

Also, others use ABC, where they developed a model for diagnosing cavernous tuberculosis and miliary tuberculosis based on chest CT scans. The first stage of this model is the pre-treatment of the CT images, which removes unnecessary distortions using a Gaussian filter. Then we move to the segmentation stage that works on pixel classification and extracting lung tissue using a region-based active contour model, and then we move to the stage of extracting the region of interest (ROI) and then extracting the features from the ROI using a region-growing algorithm. selecting the optimal subset of features using I-ABCO and SVM, and the last stage is that the features selected are used to train the radial basis function neural network classifier. They used two datasets: the first is a TB dataset consisting of 57 slides with miliary tuberculosis, 350 slides with cavernous tuberculosis, and 152 normal lung slides. The other LIDC-IDRI dataset consisted of 1018 cases of lung computed tomography scans. The search process has been improved using two evaluation methods: mutual information (MI) and rough dependency measure (RDM). where I-ABCO+MI achieved an accuracy of 88.34% for the TB dataset and 92.63% for LIDC-IDRI, and I-ABCO+RDM achieved an accuracy of 87.32% for the TB dataset and 90.17% for LIDC-IDRI. In terms of future suggestions, this work can be extended through the advancement of a CAD system and experimentation with different segmentation algorithms [33].

TABLE 1: Accuracy measured for different disease using PSO, GWO, and ABC.

Swarm intelligence algorithm	Disease name	Samples number	Accuracy	references
SPO	skin cancer	3297 images	89.17%	[22]
	Adrenal-gland tumors	120 patients	81.67%	[23]
	diabetes	314 cases	98.73%	[24]
	Heart	303 people	97.3%	[25]
		4238 people	96.1%	[26]
GWO	Diabetes		81%	[27]

	rheumatoid arthritis			86.36 %		[28]
	Parkinson	HandPD	632 cases	94.83%		[29]
		SpeechPD	195 cases			
		VoicePD	1040 cases			
	Alzheimer	287 cases		96.23%		[30]
ABC	Parkinson			PD speech 2	94.05%	[31]
				HandPD spiral	97.62%	
				PD acoustic	100%	
				PD speech	100%	
				HandPD meander	96.83%	
	Malaria			96%		[32]
	stroke	43,400 cases		87.09%		[33]
	Cavernous, miliary tuberculosis	559 slides		I-ABCO+MI 88.34%	I-ABCO+RDM 87.32%	[34]
		1018 cases		I-ABCO+MI 92.63%	I-ABCO+RDM 90.17%	

V. Conclusions

The area of automated illness diagnosis is discussed in this survey. It falls within the category of bioinformatics research. The relevance of swarm intelligence algorithms (PSO, GWO, and ABC) for categorization is concluded from this survey. It gave a thorough examination of PSO, GWO, and ABC and how they're used to diagnose various ailments. The researchers employed PSO to identify diseases such as skin cancer, adrenal gland tumours, diabetes, and coronary heart disease, and they used a variety of datasets and classification approaches and proposed future trends, as indicated above, to obtain the accuracy shown in Table 1. They also employed GWO to identify rheumatoid arthritis, Parkinson's disease, Alzheimer's disease, and diabetes, as well as ABC to detect Parkinson's, malaria, stroke, TB, cavernous tuberculosis, and miliaria tuberculosis. Although there are modern swarm intelligence algorithms such as the naked mole-rat algorithm and supply-demand-based optimisation, they have not been addressed by researchers in the field of optimisation, especially in the field of diagnosing diseases. It is preferable to go towards these algorithms and use them for diagnosing diseases in the future.

References

[1] C. B. C. Latha and S. C. Jeeva, “Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques,” *Informatics Med. Unlocked*, vol. 16, 2019, doi: 10.1016/j.imu.2019.100203.

[2] M. abdulameer kadhim alsaeedi and S. Kurnaz, “Feature selection for diagnose coronavirus (COVID-19) disease by neural network and Caledonian crow learning algorithm,” *Appl. Nanosci.*, no. 0123456789, 2022, doi: 10.1007/s13204-021-02159-x.

[3] A. E. Mohamed, “Comparative Study of Four Supervised Machine Learning Techniques for Classification,” *Int. J. Appl. Sci. Technol.*, vol. 7, no. 2, pp. 5–18, 2017.

[4] M. T. Scholar, A. Urgiriye, and A. R. Bhartiya, “Review of Machine Learning Algorithm on Cancer DataSet,” vol. 6, no. 6, pp. 3259–3267, 2020.

[5] K. P. Linthicum, K. M. Schafer, and J. D. Ribeiro, “Machine learning in suicide science: Applications and ethics,” *Behav. Sci. Law*, vol. 37, no. 3, pp. 214–222, 2019, doi: 10.1002/bsl.2392.

[6] P. Agrawal, H. F. Abutarboush, T. Ganesh, and A. W. Mohamed, “Metaheuristic Algorithms on Feature Selection: A Survey of One Decade of Research (2009-2019),” *IEEE Access*, vol. 9, pp. 26766–26791, 2021, doi: 10.1109/ACCESS.2021.3056407.

[7] J. Miao and L. Niu, “A Survey on Feature Selection,” *Procedia Comput. Sci.*, vol. 91, no. Itqm, pp. 919–926, 2016, doi: 10.1016/j.procs.2016.07.111.

[8] R. K. Sivagaminathan and S. Ramakrishnan, “A Hybrid Approach for Effective Feature Selection using Neural Networks and Artificial Bee Colony Optimization,” *Expert Syst. Appl.*, vol. 33, no. 1, pp. 49–60, 2007, doi: 10.1016/j.eswa.2006.04.010.

[9] S. M. Almufti, “Historical survey on metaheuristics algorithms,” *Int. J. Sci. World*, vol. 7, no. 1, p. 1, 2019,

doi: 10.14419/ijsw.v7i1.29497.

[10] W. Zhao, L. Wang, and Z. Zhang, "Supply-Demand-Based Optimization : A Novel Economics-Inspired Algorithm for Global Optimization," no. 1, pp. 1–16.

[11] D. Karaboga, "An idea based on honey bee swarm for numerical optimization. Technical Report TR06, Erciyes University," *Erciyes Univ.*, vol. 46(2), no. TR06, pp. 1–10, 2005, [Online]. Available: http://dx.doi.org/10.1016/j.cirp.2016.06.001%0Ahttp://dx.doi.org/10.1016/j.powtec.2016.12.055%0Ahttps://doi.org/10.1016/j.ijfatigue.2019.02.006%0Ahttps://doi.org/10.1016/j.matlet.2019.04.024%0Ahttps://doi.org/10.1016/j.matlet.2019.127252%0Ahttp://dx.doi.o

[12] J. C. Bansal, H. Sharma, S. S. Jadon, and M. Clerc, "Spider Monkey Optimization algorithm for numerical optimization," *Memetic Comput.*, vol. 6, no. 1, pp. 31–47, 2014, doi: 10.1007/s12293-013-0128-0.

[13] D. Karaboga, B. Gorkemli, C. Ozturk, and N. Karaboga, "A comprehensive survey: Artificial bee colony (ABC) algorithm and applications," *Artif. Intell. Rev.*, vol. 42, no. 1, pp. 21–57, 2014, doi: 10.1007/s10462-012-9328-0.

[14] IEEE Staff, "Particle swarm optimization algorithm and its parameters: A review," no. Iccccm, 2016.

[15] Y. Zhang, S. Wang, and G. Ji, "A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications," *Math. Probl. Eng.*, vol. 2015, 2015, doi: 10.1155/2015/931256.

[16] H. Fan, "A modification to particle swarm optimization algorithm," *Eng. Comput. (Swansea, Wales)*, vol. 19, no. 7–8, pp. 970–989, 2002, doi: 10.1108/02644400210450378.

[17] D. Venter, "Particle Swarm Optimization," *Soc. Transit.*, vol. 34, no. 2, pp. 197–205, 2003, doi: 10.1080/21528586.2003.10419092.

[18] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview," *Soft Comput.*, vol. 22, no. 2, pp. 387–408, 2018, doi: 10.1007/s00500-016-2474-6.

[19] H. Wang *et al.*, "Artificial bee colony algorithm based on knowledge fusion," *Complex Intell. Syst.*, vol. 7, no. 3, pp. 1139–1152, 2021, doi: 10.1007/s40747-020-00171-2.

[20] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[21] H. Faris, I. Aljarah, M. A. Al-Betar, and S. Mirjalili, "Grey wolf optimizer: a review of recent variants and applications," *Neural Comput. Appl.*, vol. 30, no. 2, pp. 413–435, 2018, doi: 10.1007/s00521-017-3272-5.

[22] E. B. A. Ş. Aran, "Skin Cancer Diagnosis Using CNN Features with Genetic Algorithm and Particle Swarm Optimization Methods," pp. 1–17, 2022.

[23] J. Xu, F. Tian, L. Wang, and Z. Miao, "Binary Particle Swarm Optimization Intelligent Feature Optimization Algorithm-Based Magnetic Resonance Image in the Diagnosis of Adrenal Tumor," *Contrast Media Mol. Imaging*, vol. 2022, 2022, doi: 10.1155/2022/5143757.

[24] H. Qteat and M. Awad, "Using Hybrid Model of Particle Swarm Optimization and Multi-Layer Perceptron Neural Networks for Classification of Diabetes," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 3, pp. 11–22, 2021, doi: 10.22266/ijies2021.0630.02.

[25] I. D. Mienye and Y. Sun, "Improved heart disease prediction using particle swarm optimization based stacked sparse autoencoder," *Electron.*, vol. 10, no. 19, 2021, doi: 10.3390/electronics10192347.

[26] S. S. G. and M. K., "Diagnosis of diabetes diseases using optimized fuzzy rule set by grey wolf optimization," *Pattern Recognit. Lett.*, vol. 125, pp. 432–438, 2019, doi: 10.1016/j.patrec.2019.06.005.

[27] S. Sundaramurthy and P. Jayavel, "A hybrid Grey Wolf Optimization and Particle Swarm Optimization with C4.5 approach for prediction of Rheumatoid Arthritis," *Appl. Soft Comput. J.*, vol. 94, p. 106500, 2020, doi: 10.1016/j.asoc.2020.106500.

[28] P. Sharma, S. Sundaram, M. Sharma, A. Sharma, and D. Gupta, "Diagnosis of Parkinson's disease using modified grey wolf optimization," *Cogn. Syst. Res.*, vol. 54, pp. 100–115, 2019, doi: 10.1016/j.cogsys.2018.12.002.

[29] S. K, L. S.K., A. Khanna, S. Tanwar, J. J. P. C. Rodrigues, and N. R. Roy, "Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier," *Comput. Electr. Eng.*, vol. 77, pp. 230–243, 2019, doi: 10.1016/j.compeleceng.2019.06.001.

[30] H. Li *et al.*, "A Hybrid Feature Selection Algorithm Based on a Discrete Artificial Bee Colony for Parkinson's Diagnosis," *ACM Trans. Internet Technol.*, vol. 21, no. 3, 2021, doi: 10.1145/3397161.

[31] E. Obinna and L. Kabari, "A Modified Approach to Malaria Diagnosis Using Artificial Bee Colony Algorithm," *Ijisrt.Com*, vol. 6, no. 6, 2021, [Online]. Available: https://ijisrt.com/assets/upload/files/IJISRT21JUN1103.pdf

[32] A. Dev and S. K. Malik, "Artificial bee colony optimized deep neural network model for handling imbalanced stroke data: ABC-DNN for prediction of stroke," *Int. J. E-Health Med. Commun.*, vol. 12, no. 5, pp. 67–83, 2021, doi: 10.4018/IJEHMC.20210901.oa5.

[33] A. Isaac, H. Khanna Nehemiah, and A. Kannan, "Computer-Aided Diagnosis System for Diagnosis of Cavitary and Miliary Tuberculosis Using Improved Artificial Bee Colony Optimization," *IETE J. Res.*, 2021, doi: 10.1080/03772063.2021.1946440.