T.C.
ISTANBUL SABAHATTIN ZAIM UNIVERSITY
GRADUATE EDUCATION INSTITUTE
COMPUTER SCIENCE AND ENGINEERING

A NEW HYBRID METAHEURISTIC ALGORITHM APPLIED ON A CLASSIFICATION PROBLEM

MASTER THESIS

Armir KAÇABETI

Istanbul
August-2021

T.C.

ISTANBUL SABAHATTIN ZAIM UNIVERSITY

GRADUATE EDUCATION INSTITUTE

COMPUTER SCIENCE AND ENGINEERING

A NEW HYBRID METAHEURISTIC ALGORITHM APPLIED ON A CLASSIFICATION PROBLEM

MASTER THESIS

Armir KAÇABETI

Supervisor
Asst. Prof. Dr. Amir SEYYEDABBASI

Istanbul
August-2021

APPROVAL PAGE

This study has been approved in partial fulfillment of the requirements for MA Degree in Computer Science and Engineering.

Chairperson of jury: Asst. Prof. Dr. Amir SEYYEDABBASI

Member of jury: Asst. Prof. Dr. Aydın Tarık ZENGİN

Member of jury: Asst. Prof. Dr. Ali HAMİTOĞLU

Approval by

Prof. Dr. Ali GÜNEŞ

Director, Graduate Education Institute

DECLARATION OF SCIENTIFIC ETHICS AND ORIGINALITY

This is to certify that this MA thesis dissertation titled "**A New Hybrid Metaheuristic Algorithm Applied on a Classification Problem**" is my own work and I have acted according to scientific ethics and academic rules while producing it. I have collected and used all information and data according to scientific ethics and guidelines on thesis writing of Sabahattin Zaim University. I have fully referenced, in both the text and bibliography, all direct and indirect quotations and all sources I have used in this work.

Armir KAÇABETI

Istanbul, August 2021

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to my supervisor, Asst. Prof. Dr. Amir SEYYEDABBASI, for his continuous guidance and support throughout this thesis.

Secondly, I would like to express my thanks to my friends and also DATAPROM team for their supportive and thoughtful behaviour during this process.

Last but not least, my greatest thanks goes to my beloved family, my mother, my father and especially my dear sisters for their endless support, encouragement, and motivation throughout education life.

Armir KAÇABETI

Istanbul, 2021

ABSTRACT

A NEW HYBRID METAHEURISTIC ALGORITHM APPLIED ON A CLASSIFICATION PROBLEM

Armir KAÇABETI
Master, Computer Science and Engineering
Thesis Advisor: Asst. Prof. Dr. Amir SEYYEDABBASI

August-2021, Page: 56

Working with datasets has become a very complex task nowadays due to the large number of data which are formed. Data science and data mining are two of many fields that are affected by the rapid growth of data and high data dimensions. Large data causes problems like high computational costs, memory costs and poor execution accuracy in these fields. The above mentioned problems occurs because datasets are mainly machine learning classifiers; however, the accuracy of machine learning can be affected by noise and irrelevant functions too. Feature selection has been presented as one of the solutions of these problems. Feature selection reduce the dimensionality of the data by selecting the best subset of features. This reduction helps in speeding up the learning process, simplify the learnt model and in the same time increases the performance. During this process the accuracy performance should be maintained too, which is the main aim of the feature reduction problem. In this thesis study, a new feature reduction approach has been invented to improve and approximate the solutions of the feature selection problems. To achieve this goal a new hybrid metaheuristic algorithm is proposed. In the hybridization, three well-known metaheurstic algorithms were used, while performance of each of them is significant. Exactly, Salp Swarm Algorithm, Grey Wolf Optimization and Whale Optimization Algorithm were used. The hybrid algorithm follows the main structure of these algorithm since each algorithm has specific nature behavior. Two main improvement have been included in this study. The first is improving the population diversity in the search space by using the three of the algorithms, and the second improvement is made in the exploitation phase by mostly using GWO and WOA.

Keywords: Salp Swarm Algorithm, Grey Wolf Optimization, Whale Optimization Algorithm, Feature Selection, Meta-heuristic Algorithms, Swarm Intelligence

ÖZET

BİR SINIFLANDIRMA PROBLEMİNE UYGULANAN YENİ BİR HİBRİT METASEZGİSEL ALGORİTMA

Armir KAÇABETİ
Yüksek Lisans, Bilgisayar Bilimi ve Mühendisliği
Tez Danışmanı: Dr. Öğr. Üyesi Amir SEYYEDABBASİ

Ağustos-2021, 56 Sayfa

Veri kümeleri ile çalışmak, verilerin çoğalmasıyla nedeniyle günümüzde çok karmaşık bir görev haline gelmiştir. Veri bilimi ve veri madenciliği, verinin hızlı büyümesinden ve yüksek veri boyutlarından etkilenen birçok alandan ikisidir. Büyük veri, bu alanlarda yüksek hesaplama maliyetleri, bellek maliyetleri ve zayıf yürütme doğruluğu gibi sorunlara neden olur. Yukarıda bahsedilen sorunlar, veri kümelerinin temel olarak makine öğrenimi sınıflandırıcıları olması nedeniyle oluşur; ancak makine öğreniminin doğruluğu gürültüden ve alakasız işlevlerden de etkilenebilir. Bu problemlerin çözümlerinden biri olarak öznitelik seçimi sunulmuştur. Özellik seçimi, özelliklerin en iyi alt kümesini seçerek verilerin boyutsallığını azaltır. Bu azalma, öğrenme sürecini hızlandırmaya, öğrenilen modeli basitleştirmeye ve aynı zamanda performansı artırmaya yardımcı olur. Bu işlem sırasında, özellik azaltma probleminin temel amacı olan doğruluk performansı da korunmalıdır. Bu tez çalışmasında, öznitelik seçimi problemlerinin çözümlerini iyileştirmek ve yaklaşık çözümleri tahmin etmek için yeni bir öznitelik indirgeme yaklaşımı icat edilmiştir. Bu amaca ulaşmak için yeni bir hibrit metasezgisel algoritma önerilmiştir. Hibridizasyonda, her birinin performansı önemli olmakla birlikte, iyi bilinen üç metasezgisel algoritma kullanılmıştır. Salp Sürü Algoritması, Gri Kurt Optimizasyonu ve Balina Optimizasyon Algoritması kullanıldı. Hibrit algoritma, her algoritmanın kendine özgü doğa davranışı olduğundan, bu algoritmaların ana yapısını takip eder. Bu çalışmaya iki ana iyileştirme dahil edilmiştir. Birincisi, algoritmaların üçünü kullanarak arama uzayındaki popülasyon çeşitliliğini iyileştirmek, ikincisi ise daha çok GWO ve WOA kullanılarak sömürü aşamasında yapılır.

Anahtar Kelimeler: Salp Sürü Algoritması, Gri Kurt Optimizasyonu, Balina Optimizasyon Algoritması, Özellik Seçimi, Meta-sezgisel Algoritmalar, Sürü Zekâsı

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Nowadays the real-global issues usually consist of a massive range of records which shape datasets. Handling this records has end up a complicated task. As we know, a dataset consists of a massive range of attributes/features. Most of the time not all the features contain useful information for the datasets. Irrelevance and redundancy take part in some of those features, which causes the degradence of the performance of the model, reduce the learning speed and increase the computational complexity of built models. This is a major problem in the field of data mining, knowledge discovery, and machine learning. Feature reduction one of the solutions of this problem. It selects a subset of salient features, eliminates irrelevant and redundant features, and reduce the dimensionality of the dataset. This reduction helps to accelerate the learning process, simplify the learnt model and improve the performance. While reducing the size of datasets we have to maintain the accuracy performance too, which is the principle goal of feature reduction problem.

Feature reduction can be made in two ways: by feature construction or by feature selection. The feature construction creates a new feature set from original datasets. Feature selection is composed of choosing the relevant features from the original dataset.

Feature selection based on the subset evaluation process can be grouped in two major categories which are: filter-based method and the wrapper-based method. In the filter-based methods, features are being evaluated without considering any classification algorithms. Firstly, features rank by using measures such as distance, dependency and consistency and then the model is created by learning on train data with the selected features. Wrapper is the opposite of the filter method as it utilizes a learning algorithm to evaluate a candidate feature subset. The wrapper method is more accurate than the first method, but it is computationally more expensive.

Feature selection is in fact a binary optimization problem, in which solutions are bounded to 0 or 1 as binary values. Every solution in this problem is represented in a binary vector.

The number of the features that exists within the dataset is represented by the length of the vector. If the value is one means the feature is selected; otherwise, the feature is not selected. Using a stochastic search approach like metaheuristic algorithms to explore a large portion of the search space and exploit near-optimal solutions is a way of tackling feature selection. In recent years, metaheuristic algorithms due to their ability in solving different problems in various fields have attracted the eye of the many researchers. Accordingly, there have been proposed many metaheuristic algorithms inspired from almost different phenomena in the nature, math, or physics. Population-based metaheuristic algorithms applied on feature selection had shown improved results. For instance, ant colony optimization (ACO), which mimics the foraging behavior of ants, has been employed as a wrapper feature selection method. Bat algorithm (BA) is another algorithm that has been successfully used in feature selection problem. Gravitational search algorithm (GSA), based on law of motion and gravitation (invented by Newton), has been employed as a wrapper feature selection too. Whale Optimization Algorithm (WOA) is another recent algorithm that mimics the social behavior of humpback whales was used as a wrapper feature selection method. Dragonfly algorithm (DA) is another recent metaheuristic algorithm has been applied on feature selection problem.

In this thesis, a feature selection approach is proposed. The approach is evaluated on hybridizing three metaheuristic search algorithms, Salp Swarm Algorithm (SSA), Gray Wolf Optimizer (GWO) and Whale Optimization Algorithm (WOA). The proposed new hybridized algorithm is called SSAGWOWOA. The process starts with the SSA and in the other steps, based on the conditions of the iterations and random values the other two algorithms are executed too. The purpose is to improve the SSA in the phases of exploration and exploitation and then to use the new approach for the feature selection, exactly for image feature reduction.

Chapter 2 is focused in giving an overview of metaheuristics, also describing in details the algorithms that are hybridized in this thesis. The 3rd chapter introduces the proposed approach while in the 4th chapter simulation and results are explained.

CHAPTER II

LITERATURE REVIEW

2.1 Metaheuristic Algorithms

In computer science exists some problems that are difficult to solve by classic methods and why not sometimes impossible to solve, which are known as NP problems. To solve those problems heuristic techniques had been invented. A heuristic, in short, is a technique designed for solving NP problems, or for finding an approximate/acceptable solution when the exact solution is not founded by classic methods. Meanwhile, a metaheuristic is an iterative approach that guides a subordinate heuristic by combining intelligently different concepts for exploring and exploiting search space. They are mostly inspired by the phenomenas that occure in nature. We can define a metaheuristic also as a higher-level procedure or heuristic created to find or select a heuristic that may offer an approximate solution to an optimization problem. Metaheuristics make few assumptions while solving the optimization problems, which make them usable for a variety of problems.

Metaheuristic techniques have gain popularity over the past two decades and have been applied in different branches of science and industry. This popularity is due to the below reasons:

a) **Simplicity**: Most of them have been inspired by simple concepts which are related to physical phenomenas, animals' behaviors, or evolutionary concepts. It is precisely the simplicity that allows scientists and researchers to simulate various natural concepts, improve current meta-heuristics, hybridize two or more of them or recommend new meta-heuristics.

b) **Flexibility**: The meta-heuristics are easily applicable to various problems without making any particular changes in it's structure. They also take issues (problems) as black boxes. This means that the most important thing of a system for a meta-heuristic is the input and output.

c) **Derivation-Free (Gradient-Free) Mechanism**: The problems are optimized stochastically in metaheuristics, which means that the optimization begins with random solutions, and for finding the optimum there is no need to calculate the derivative of search spaces.

d) **Local Optima Avoidance**: Meta-heuristics have better capabilities to avoid local optima. The stagnation in local solutions is avoided and the entire search space is being searched due to the stochastic nature of metaheuristics. So meta-heuristics seems to be a good solution for optimizing problems that have a massive number of local optima.

Meta-heuristic algorithms can offer reasonable solutions within an appropriate time. Just as from time to time meta-heuristic algorithms won't assure the best solutions, additionally from time to time the solution founded won't be acceptable. In other words, the overall performance of the evolved algorithms may differ based in the problem that they are applied. Just as an algorithm may be very successful in the solution of a problem, in the other hand at the same time it may not be a good solution for another problem. As said above, the meta-heuristic algorithm is independent from the problem and is executed based on random inputs and acquired outputs. The purpose is to have an efficient algorithm which is able to make good and acceptable solutions.

2.1.1 Characteristcs of Metaheuristic Algorithms

The two main classes of Meta-heuristics are:

- Single Solution Based
- Population Based

In the first class the search process begins with one candidate solution and then is improved during the iterations, while Population-based meta-heuristics make the optimization using a set of solutions (population). In contrast with single solution base, in population based the search process begins with a random population (a couple of solutions), which is improved through iterations.

Even though meta-heuristics differs from each other, they have a common feature too. The search process of both is divided into two phases: **exploration** and **exploitation**. **The exploration** phase consists of searching the promising area(s) of the search space as broadly as possible. This phase is supported by stochastic operators of algorithms which (stochastic operators) randomly and globally search the search space. **The exploitation** consists of the local search capability around the promising regions obtained in the exploration phase.

The performance of an algorithm is specifically influenced by a fine balance among these two phases. Also, balancing these two phases is taken into consideration as a tough task. If the exploration is too little and the exploitation is too much the system may be trapped in local optima. In this case, it'd be very hard or perhaps not possible to find the global optimum.

2.1.2 Classification of Metaheuristic Algorithms

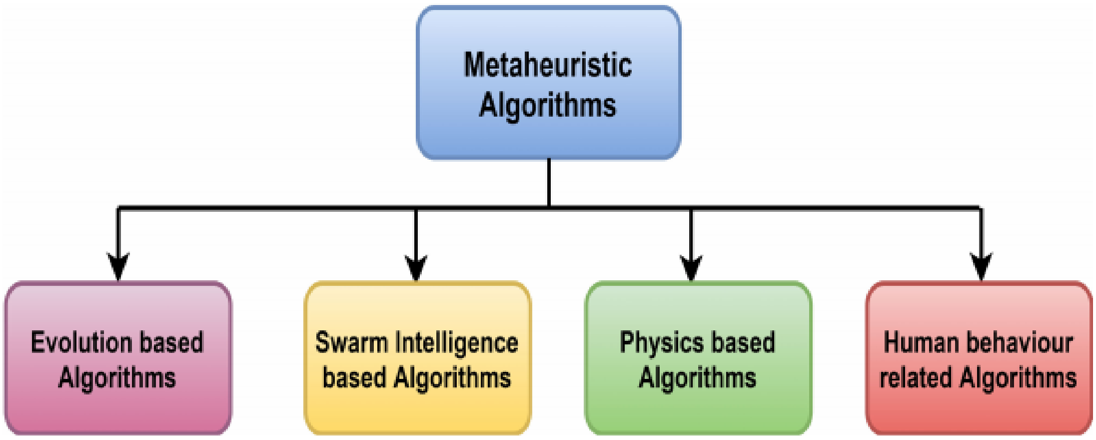

Figure 1 Classification of Metaheuristic Algorithms

According to their behavior, the metaheuristic algorithms can be grouped in 4 categories:

i. Evolution-Based

As may be understood by the name, this group is inspired from the natural evolution. To begin their process, they fisrtly generate random population of solutions. In those kind of algorithms, the best solutions come together to create new individuals, which are formed by using mutation, crossover and pick out the best solution. The most well known algorithm of this category is with no doubt, the algorithm inspired by Darwin Evolution, Genetic algorithm (GA). We can mention here other algorithms like Evolution Strategy (ES), Tabu Search (TS), Genetic Programming (GP), Differential Evolution (DE) etc.

ii. Swarm Intelligence-Based

Social behaviors of creatures like insects, animals, fish or birds etc. have been the main inspirer for this group of algorithms. Particle Swarm Optimization (PSO) is the most famous technique which is developed by Kennedy and Eberhart. In this case the behavior of a group of birds have been the inspiration of the authors. They can find their best location by flying acros the search space. We can mention here other algorithms like: Ant Colony Optimization (ACO), Monkey Optimization etc.

iii. Physics-Based

The rules of physics in the universe have also been an inspiration to scientists which for inventing the physic based group algorithms. Harmony Search (HS), Simulated Annealing (SA), etc. are examples of mentioned group.

iv. Human-Related

The last group of algorithms are inspired by human behaviors. Researchers have developed the algorithms by analyzing the human being activities, always keeping in mind that every human being (person) has its manner of doing activities, which has its affect in performance. Some of the famous algorithms are: Teaching Learning-Based Optimization Algorithm (TLBO), League Championship Algorithm etc.

2.1.3 Related Work

The majority of the metaheuristics have been developed before and in the early of the 2000’s. We can mention here three major metaheuristic algorithms: Genetic Algorithm (GA) (Goldberg, 1989), Particle Swarm Optimization (PSO) (Wei and Qiqiang, 2004) and Ant Colony Optimization (ACO) (Dorigo and Birattari, 2010).

Genetic algorithm is the most popular algorithm. Evolution theory of Charles Darwin was the main impact in developing this algorithm. GA consists of generation reproduction which includes phases like selection, crossover and mutation. It has been successfully applied in fields like soft computing, health sciences and civil engineering.

Inspired by social behavior of birds, Particle Swarm Optimization (PSO) is another populary algorithm in this category. In PSO particles communicate between each other, which move on search space. The best solution is determined by fitness function. Many of optimization problems has been solved by solutions that PSO can provide. It is mostly used as a solution in the best route-finding issues. Finding routes in these kind of problems seems to be an esssential issue especially in case of changing paths frequently. Other parameters such as power, delay, delivery rate etc. have their effect in the decision mechanism too. In those class issues, PSO can construct the most beneficial solutions and routing paths.

Ant Colony Optimization (ACO) is the perfect example of swarm intelligence algorithms. It simulates behaviors of ants as they find their way in foraging. ACO influences some factors that each ant needs to find path according to the experience of other ants in the path. The most used field of ACO is routing in wireless sensor networks. Finding the optimal route is a critical issue in these networks.

Maybe not as known as the previous algorithms, the below algorithms have been developed too:

Genetic Programming (GP) (Banzhaf et al., 1998), Differential Evolution (DE) (Storn and Price, 1997), Simulated Annealing (SA) (Van Laarhoven and Aarts, 1987), Tabu Search (TS) (Glover and Laguna, 1998), Greedy Randomized Adaptive Search Procedure (GRASP) (Marques-Silva and Sakallah, 1999), Artificial Immune Algorithm (AIA) (Dasgupta, 2012), Iterated Local Search (ILS) (Louren¸co et al., 2003), Chaos Optimization Method (COM) (Li and Jiang, 1997), Scatter Search (SS) (Mart´ı et al. in 2006), Shuffled Frog-Leaping Algorithm (SFLA) (Eusuff and Lansey, 2003), Variable Neighborhood Search (VNS) (Mladenovi´c and Hansen, 1997) etc.

Despite the successful achievements of the metaheuristic algorithms in the early 2000's, new and evolving approaches have also emerged successfully in the last decade. SSA, GWO and WOA metaheuristics are three of them that are described in detail below. They have been in the attention of many researchers and have been cited many times in recent years. Each of them is based on its kind swarm behaviour, all of these algorithms are modified for solving feature selection problems. For each of the algorithms, are provided details about their inspirations, mathematical models and pseudocodes.

A. SSA

Authors have been inspired by the swarming behavior of Salps. They have a complex life cycle and live in environments that are too difficult to access. Their shape is given in Fig. 2, and in the next Fig. 3 is also shown the ability of these creaters to live in chain, which illustrates the swarming behavior. We have to keep in mind that the extreme environments in which they live cannot be created in laboratory environments which make the researches even more difficult.

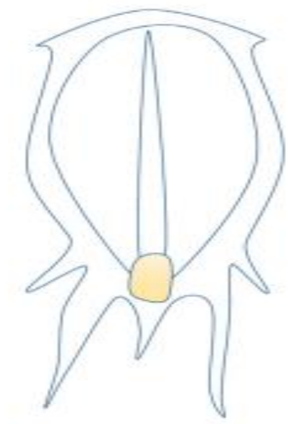

Figure 2 Shape of Salp

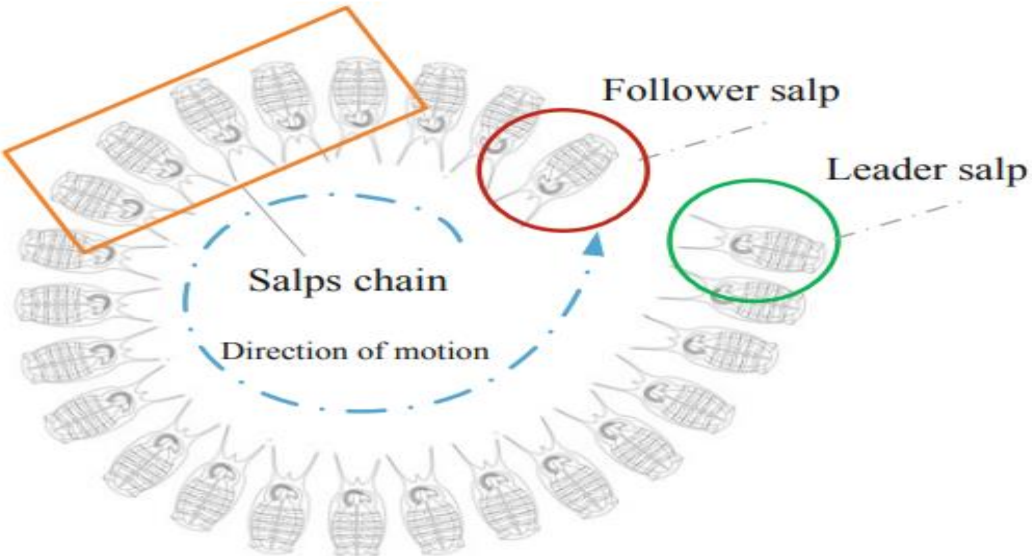

Figure 3 Salp Swarm (Salp Chain)

Mathematical Model

The population of salps consists of the leader and the followers. The first salp of the chain is considered as the leader and the rest are followers. The duty of the leader is to lead the swarm while the followers has to follow each other. In some way they indirectly follow the leader, because all the chain follows the previous salp except the leader. The mathematical model of the chain is developed by defining the positions of the salps in an n-dimensional search space (n represents the number of the variables of the given problem). So in this model too, as in the other swarm-based techniques, the position of all salps is stored in a two dimensional matrix which is called x. Also, the food source F is assumed as the swarms target.

The leader updates its position by the below equation (1)

$$X_j^1 = \begin{cases} X_{Best_j} + c_1\left((ub_j - lb_j)c_2 + lb_j\right) & if\ c_3 \geq 0.5 \\ X_{Best_j} - c_1\left((ub_j - lb_j)c_2 + lb_j\right) & else \end{cases} \quad (1)$$

where X_j^1 and X_{Best_j} represent the new position of the leader and food source respectively in the j^{th} dimension. The upper and lower bounds of j^{th} dimension are represented by ub_j and lb_j. c_2 and c_3 are variables that are generated randomly in the interval [0, 1]. In SSA, the balance between exploration and exploitation is controlled by the c_1 parameter. c_1 is calculated by the Eq. (2), where t is the current iteration and T is the maximum number of iterations.

$$c_1 = 2e^{-(\frac{4t}{T})^2} \quad (2)$$

The position of the followers is updated by using the Newton’s law of motion as in Eq. (3):

$$X_j^i = \frac{1}{2}gt^2 + \omega_0 t\ , i \geq 2 \quad (3)$$

where X_j^i represents the position of i^{th} follower salp in the j^{th} dimension. The time t corresponds to the current iteration, g indicate the acceleration and ω_0 indicate the velocity. In Eq.(3) having the initial speed ω_0 fixed to 0 and the discrepancy fixed to 1 ($\Delta t = 1$), means that the updating process of followers is expressed by a new equation as in Eq.(4):

$$X_j^i = \frac{1}{2}(X_j^i + X_j^{i-1}) \quad (4)$$

Pseudocode of SSA is shown in the below figure (4)

```
Initialize the population size N and max iterations number T.
Set the initial iteration number t := 0.
Generate the initial population X which contains N.
Evaluate solutions the fitness function of all individuals X.
Denote the best solution in the population as X_Best
Repeat
   Update c_1 according to Eq.(5).
  for i=1 to N do
     if (X_i leader) then
          Update the position of the leader salp as in Eq.(4).
      else
          Update the position of the follower salp as in Eq.(7).
      end if
   end for
   Set t = t + 1.
until (t<T).
Produce the best solution.
```

Figure 4 Pseudocode of SSA

As each metaheuristic algorithm the SSA has its own advantages and disadvantages too, which are mentioned as follows:

1) Easy to implement

2) During optimization process, SSA uses only the optimal solution in the actual iteration as food.

3) Simplicity

The disadvantages of SSA are as follows:

1) The leader is followed by the followers, if the leader falls to the local optimum, the entire population will be missleaded into the local optimum too.

2) The position of the leader is being updated based on the food's location. The calculating process reduces the direct interaction between leader and food.

3) The mathematical model of SSA expresses a low precision of algorithm because there is no transition between exploitation and exploration.

B. GWO

Grey Wolf take part in the Canidae family. Grey wolves are considered top predators, which means that they are at the top of the food chain. Grey wolves usually choose to live in a pack, which average size is 5–12. Their social hierarchy is shown in Fig. 5

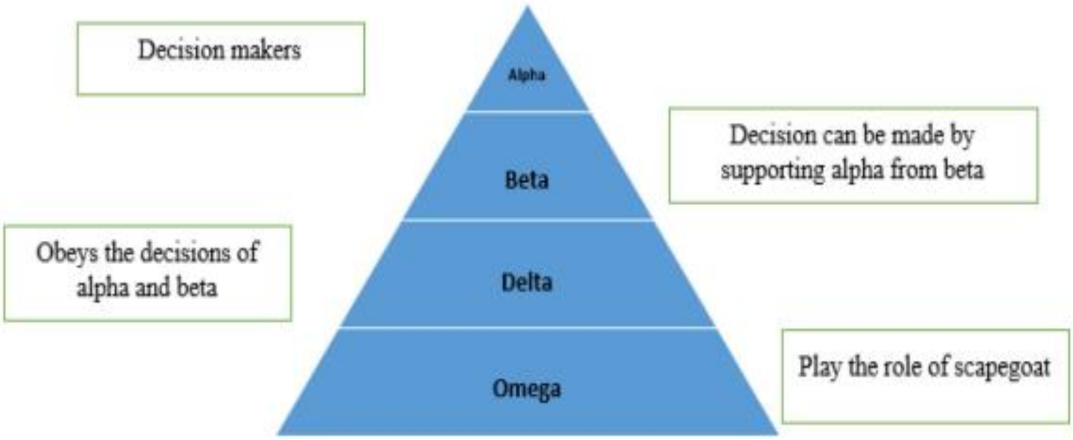

Figure 5 Hierarchy of GWO

The alpha group is also known as the leader group and is made by a male and a female wolf. The alphas have the responsibility of deciding about hunting, time to wake, sleeping place etc. Their decisions are forwarded and must be followed by the pack, which makes them to be known as dominant wolf. An interesting fact is that the alpha has to be the best in managing the pack and does not have to be the strongest member of the pack. This means that the organization of a pack is far more important than its strength.

Beta comes after the alpha group and is the second level in the hierarchy of grey wolves. They are also known as co-leaders because they assist the alpha in decision-making and other pack activities. As alphas, the betas gender is not important too, they can also be female or male. Also if one of the alpha wolves dies or “retires” the wolves of this group are the best candidate to replace the alpha. Role of this group suits to the role of an advisor for the alpha.

The third group of wolves are the delta wolves. As it can be understood by the hierarchy they dominate the 4^{th} group but have to submit to alphas and betas. This group is responsible for protecting, watching the territory, hunting, providing food and caring for the pack. The safety of the pack has to be guaranteed by this group and in case of any danger they have to warn the pack too.

The lowest level of grey wolf hierarchy is omega. They must submit to all the upper level wolves. Also, they are the last wolves that are allowed to eat.

Mathematical Model

Analyzing the hunting process of the wolves have helped the authors to develop the mathematical model. This process consists of three steps which are:

- Track, chase and approach the prey
- Pursuit, encircle and harass the prey until it stops moving
- Attack

As we mentioned above, gray wolves surround prey throughout the hunt. The following equations are proposed to mathematically model encircling behavior:

$$\vec{D} = |\vec{C}\,\vec{X_p}(t) - \vec{X}(t)| \quad (5)$$

$$\vec{X}(t+1) = \vec{X_p}(t) - \vec{A}\,\vec{D} \quad (6)$$

where t is the current iteration, $\vec{A}$, $\vec{C}$ are coefcient vectors, $\vec{X}$ is the position vector of the grey wolf, and $\overrightarrow{X_p}$ is the position vector of the prey. $\vec{A}$, $\vec{C}$, and $\vec{a}$ are calculated as Eqs. 7, 8, and 9, respectively

$$\vec{A} = 2\vec{a}\,\overrightarrow{r_1} - \vec{a} \ (7)$$

$$\vec{C} = 2\,\overrightarrow{r_2} \ (8)$$

$$\vec{a} = 2(1 - \frac{t}{T}) \ (9)$$

where $\vec{a}$ is linearly decreased from 2 to 0 through the iterations. It is used to get closer to the solution range. $\overrightarrow{r_1}$ and $\overrightarrow{r_2}$ are the random vectors in range of [0,1].

Dispite the ability of grey wolves to surround prey position, in the mathematical model the position of the prey is uncertained. Three best candidate solutions are represented by alpha (which is the frst best solution), beta, and delta. Omega wolves update their positions according to the wolves in the above layer, the delta wolves. The below Eqs. 10, 11, and 12 have been proposed:

$$\overrightarrow{D_\alpha} = |\overrightarrow{C_1}\,\overrightarrow{X_\alpha} - \vec{X}|$$

$$\overrightarrow{D_\beta} = |\overrightarrow{C_2}\,\overrightarrow{X_\beta} - \vec{X}|$$

$$\overrightarrow{D_\delta} = |\overrightarrow{C_3}\,\overrightarrow{X_\delta} - \vec{X}| \ (10)$$

$$\overrightarrow{X_1} = |\overrightarrow{X_\alpha} - \overrightarrow{A_1}\,\overrightarrow{D_\alpha}|$$

$$\overrightarrow{X_2} = |\overrightarrow{X_\beta} - \overrightarrow{A_2}\,\overrightarrow{D_\beta}|$$

$$\overrightarrow{X_3} = |\overrightarrow{X_\delta} - \overrightarrow{A_3}\,\overrightarrow{D_\delta}| \ (11)$$

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1}(t) + \overrightarrow{X_2}(t) + \overrightarrow{X_3}(t)}{3} \ (12)$$

As we already know, exploration and exploitation have critical importance in meta-heuristic algorithms. GWO attempts to trade-of among those phases. In GWO, $\vec{a}$ value in each iteration decreases from 2 to 0. $\vec{A}$ value is also decreased by $\vec{a}$. The value of $\vec{A}$ is critical to a grey wolf. When |A|<1, wolves try to search other prey, which proves the exploration and exploitation concepts. $\vec{C}$ parameter gives random values in each iteration. This value has an efect in exploration all time, even in the final iteration, authors say.

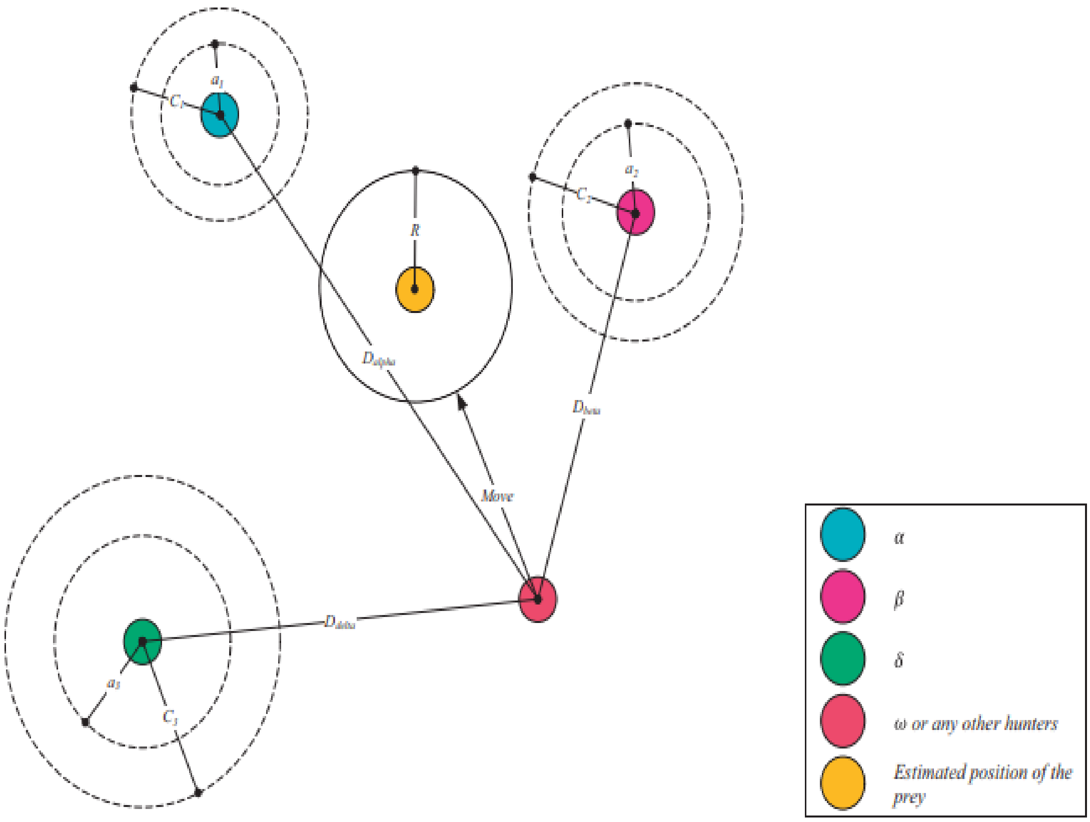

Figure 6 Position updating in GWO

Pseudocode of GWO is shown in the below figure (7)

```
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
```

Figure 7 Pseudocode of GWO

Advantages and Disadvantages

GWO shows a good performance against alternative metaheuristic algorithms, such as PSO or GSA. Compared to others, in the exploitation phase seems to be very competitive, while having good results in the exploration phase too.

Anyway, GWO has troubles almost about the stability (balance) among exploration and exploitation. Another disadvantage of it is being not able to solve nonlinear equation systems.

C. **WOA**

Has been inspired by the bubble-net hunting strategy. They have a special hunting method which consist of hunting small fishes close to the surface by creating bubbles in the shape of 9. This method has been showed in the Fig. 8.

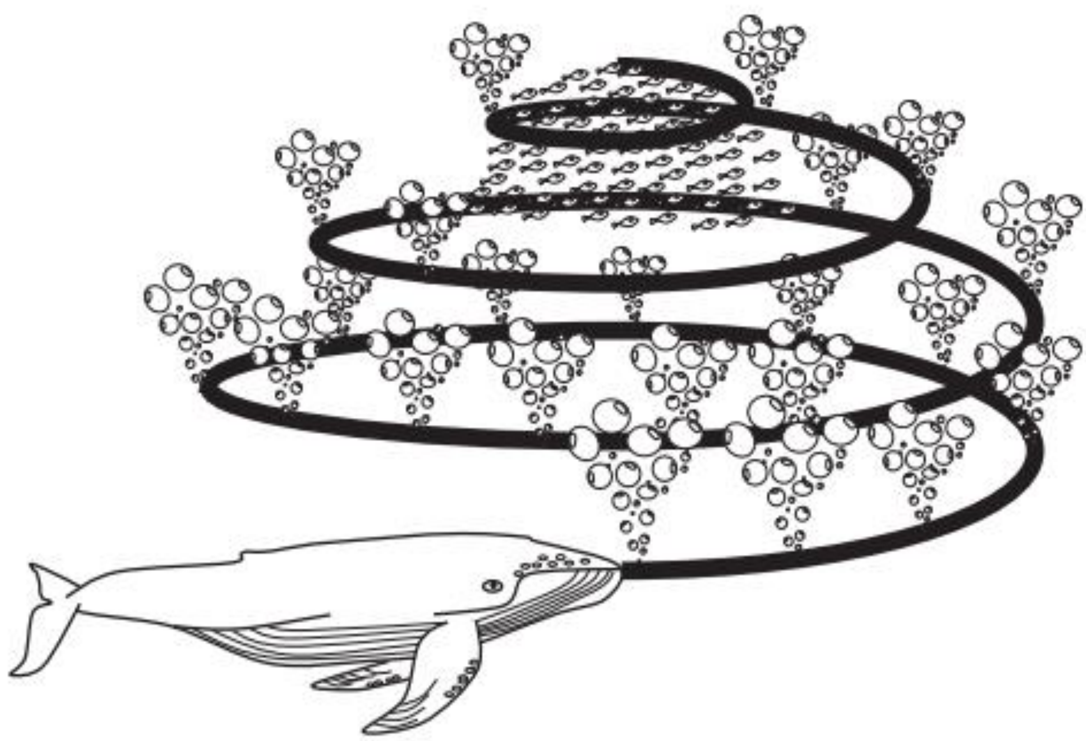

Figure 8 Bubble-net encircling strategy of WOA

In the beginning, this behavior has been investigated from surface while later some tag sensors are set to investigate more accurately this behaviour. They found two main maneuvers associated with bubble, named as 'upward-spirals' and 'double-loops'. Firstly, humpback whales dive around 12 m down and then start to create bubble in a spiral shape around the prey and then swim up toward the surface. The second maneuver consists of three different stages which are: coral loop, lobtail, and capture loop. As a result, has been discovered that they own a unique behaviour that can not be found in other whales.

Mathematical Model

Humpback whales have the potential to understand the placement of prey and so can flank them. In WOA, the best candidate solution is the target prey. When the best search agent is defined, the rest of search agents try to update their positions towards the best agents by using the following equations:

$$\vec{D} = |\vec{C}\,\overrightarrow{X^*}(t) - \vec{X}(t)| \text{ (13)}$$

$$\vec{X}(t+1) = \overrightarrow{X^*}(t) - \vec{A}\vec{D} \text{ (14)}$$

where t is the current iteration, $\vec{A}$ and $\vec{C}$ are coefficient vectors, $\overrightarrow{X^*}$ is the position vector of the best solution obtained, $\vec{X}$ is the position vector and | | is the absolute value. If a better solution is found then the $\overrightarrow{X^*}$ should be updated in each iteration. The vectors $\vec{A}$ and $\vec{C}$ are calculated as follows:

$$\vec{A} = 2\vec{a}\,\vec{r} - \vec{a} \text{ (15)}$$

$$\vec{C} = 2\,\vec{r} \text{ (16)}$$

where $\vec{a}$ is linearly decreased from 2 to 0 through iterations (in both exploration and exploitation phases) and $\vec{r}$ is a random vector which values are in [0,1]. The region of the whale in which close to the prey can be managed by values of $\vec{A}$ and $\vec{C}$ vectors. The new position of the search agent can be diagnosed among the current position of the whale and the best position by assigning values for $\vec{A}$ within the range [- 1 and 1].

The distance between the best position X_i and the current position X, is calculated by the Eq. (17) which also is being used to create a spiral-shaped approach.

$$\vec{X}(i+1) = e^{bk}\cos(2\pi k)\,\overrightarrow{D^*} + \overrightarrow{X^*}(i) \text{ (17)}$$

where the distance between the whale and prey (which is the best solution) is represented by $\overrightarrow{D^*}$.

$$\overrightarrow{D^*} = |\overrightarrow{X^*}(i) - \vec{X}(i)| \quad (18)$$

b is a constant value that identifies the logarithmic spiral shape and k represents a random number in the range [- 1 and 1]. The circular contraction mechanism and spiral mechanism have both a 50% chance of being chosen during the iterations as shown in Eq. (19).

$$\vec{X}(i+1) = \begin{cases} \overrightarrow{X^*} - \vec{A}\vec{D} & , \; if\ p < 0.5 \\ e^{bk}\cos(2\pi k)\,\overrightarrow{D^*} + \overrightarrow{X^*}(i), & if\ p \geq 0.5 \end{cases} \quad (19)$$

where p is a number between [0 and 1].

The exploration phase is a random search method. This method extends the exploratory phase, where hunting for prey depends on changing the position of each whale. At this stage, the position of the whale will change, depending on the random search, not the optimal position. This technique is resulted in performing global optima and overcoming local optima:

$$\vec{X}(i+1) = \overrightarrow{X_{rnd}} - \vec{A}\vec{D} \quad (20)$$

$$\vec{D} = |\vec{C}\,\overrightarrow{X_{rnd}} - \vec{X}| \quad (21)$$

where $\overrightarrow{X_{rnd}}$ is the randomly selected position of the whale among the whales. As can be seen in the pseudo-code of WOA, the population is being initialized randomly, and then the applicability of each search agent is evaluated. This process continues until the best solution is found. Then, the variable coefficients are updated and a random number used to update the location of the agent using equations (14) and (20) or equation (17).

Convergence is being guaranteed by WOA because the position is updated based on the best fit. Therefore, WOA can be maintained at the local optimum, and since it linearly decreases from 2 to 0, a is the main factor affecting the two-phase balance.

Pseudocode of WOA is shown in the below figure (9)

```
Initialize the whales population Xi (i = 1, 2, ..., n)
Calculate the fitness of each search agent
X*=the best search agent
while (t < maximum number of iterations)
    for each search agent
    Update a, A, C, l, and p
        if1 (p<0.5)
            if2 (|A| < 1)
                Update the position of the current search agent by the Eq. (2.1)
            else if2 (|A| ≥ 1)
                Select a random search agent (Xrand)
                Update the position of the current search agent by the Eq. (2.8)
            end if2
        else if1 (p≥0.5)
                Update the position of the current search by the Eq. (2.5)
        end if1
    end for
    Check if any search agent goes beyond the search space and amend it
    Calculate the fitness of each search agent
    Update X* if there is a better solution
    t=t+1
end while
return X*
```

Figure 9 Pseudocode of WOA

2.2 Feature Selection in Metaheuristic Algorithms

Feature selection techniques were implemented to classification problems in order to select a reduced feature set which improves the classifier by making it more accurate and faster. For a large number of features, comparing all states is computationally non-feasible requiring meta-heuristic search methods. These methods try to solve the challenges that related to real world problems with competition and cooperation strategy between agents. Many studies have been done in the feature selection, which are intersection with swarm intelligence. Advanced binary ACO (ABACO) was proposed for feature selection and dimension reduction. In 2005, Şahan et al. applied the Attribute Weighted Artificial Immune System (AWAIS) to diagnose Heart and Diabetic diseases. In this study has shown the negative effects of irrelative features in diseases diagnosis process. In other study, Huang proposed a hybrid method of ACO and SVM for feature selection. Inbarani et al. offered a model based on PSO and rough sets strategy for selected features. Nahar et al. used the computational intelligence to diagnose Heart disease.

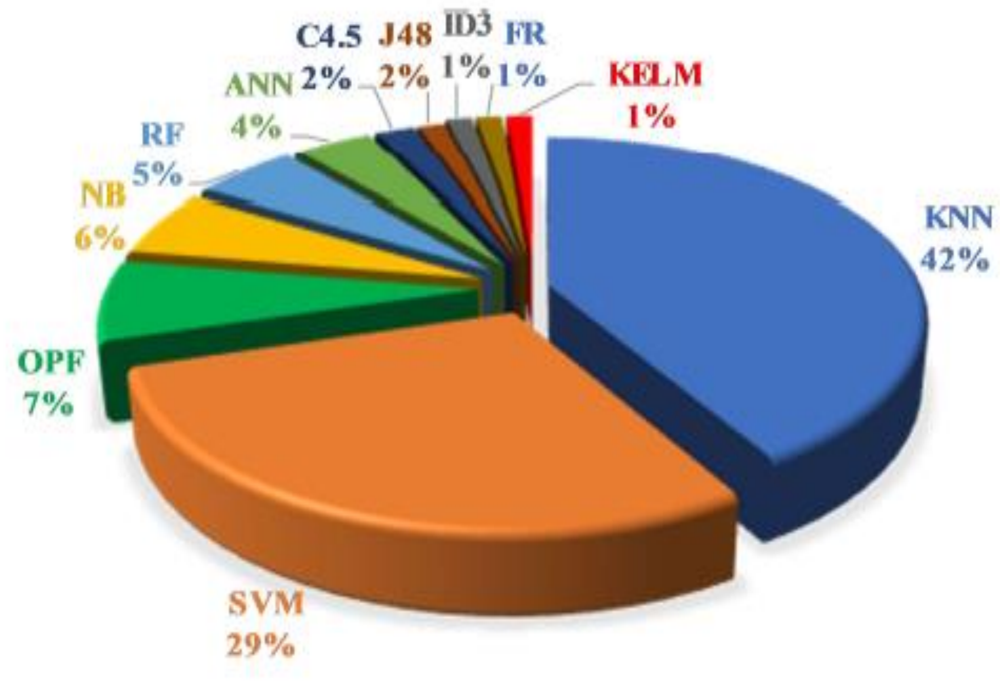

Figure 10 Role of different classifiers in solving feature select problems

2.2.1 Related Work of SSA in Feature Selection

Ibrahim used SSA to solve the feature selection problem (SSAFS) for the first time in 2017 applying a threshold of 0.5 to construct a binary vector. SSAFS is implemented to health records like breast, bladder, and colon cancer records and compared with other algorithms. SSA with ten chaotic maps and transfer functions. The KNN classifier is used to asses the classification accuracy and applied to twenty reference data sets. Faris has evolved an efficient binary SSA with S-shaped and V-shaped transfer functions, which incorporates a crossover operator in place of an intermediate operator to enhance scanning quality. The proposed method is used with the KNN classifier and applied to 22 known ICU data sets, and it is found that the S-shaped transfer function gives the best results. In order to remove the traps in the local optimization and improve the exploration and utilization of SSA, the Salpa position is updated with the chaotic singer map and local tubisat search algorithm. Applied to twenty reference records and three hadith records. Hegazy improves SSA (ISSA) by adding weights to match the best solution for KNN classifier rendering and classification. ISSA was used for 23 ICU records and compared with baseline SSA and four other meta-heuristic algorithms.

2.2.2 Related Work of GWO in Feature Selection

In the first binary version of GWO, Amari used the sigmoid transfer function to obtain the binary vector (bGWO). The KNN classifier was applied to 18 different ICU data sets in order to calculate the classification accuracy. In addition, for good research, small, random and large initialization methods are used in the initialization phase. A modified version of GWO was proposed by Sharma for identifying symptoms of Parkinson's disease using random forest, KNN, and decision tree classifiers. Patak proposed to charge GWO. Flight is used to select the appropriate attributes of the Random Forest A classifier to analyze the image strategy and apply it to the Bossbase version 1.01 data set. The results obtained show excellent performance to achieve excellent convergence. To diagnose cardiovascular diseases, AlTashi uses the GWO algorithm to select the best traits and uses SVM as a classifier. The proposed method has been applied to the

Cleveland dataset, which is free to use and has very good performance. In addition, the author proposes a binary version (BMOGWS), which uses a sigmoid function to solve the multi-criteria feature selection problem using artificial neural networks for classification. BMOGWS was applied to fifteen reference data sets and compared with MOGWO with tanh transfer function. Hu proposed a new transfer function and a new scheme to update GWO parameters. Advanced GWO (ABGWO) was applied to 12 ICU data sets and showed better results than other algorithms. GWO has different versions for classification in different areas such as medical diagnosis, cervical cancer, electromyography (EMG) signals, facial emotion detection, and text function selection.

2.2.3 Related Work of WOA in Feature Selection

Hussien used S-shaped and V-shaped transfer functions in conventional WOA in order to solve the binary optimization problem, and in 2017 used 11 UCI data sets to solve the feature selection problem. Their relevance. bWOA has proven that it can provide maximum accuracy by minimum selected features. In order to improve the proposed method, Syed proposed a chaotic whale optimization algorithm (CWOA) with ten chaotic cards. The best balance between the two important attributes of the study and the use of the algorithm. Tubishat classifies the Arabic sentiment analysis dataset by proposing an improved WOA (IWOA). IWOA contains evolutionary operators, such as crossover, mutation, and selection, just like in differential evolution. Data sets are available and compared with other methods. Jalili proposed two binary variants of WOA, including crossover and mutation operators and the use of roulette and tournament selection in WOA. 20 records are used in this method. A new version of WOA was proposed by Agrawal based on the quantum concept, using a qubit representation that is applicable to everyone, and applying the new version to fourteen data sets. There are several other versions of WOA to solve the problem of feature selection.

2.3 Hybrid Metaheuristic Algorithms

Hybrid meta-heuristic algorithm refers to combining the best operators in different meta-heuristic algorithms to develop a new and improved algorithm. In recent years, hybrid algorithms have received a lot of attention in solving optimization problems. Developed to obtain a subset of relevant and best features from the original data set. There are many ways to build better algorithms to find the best solution. The improved algorithm helps to eliminate the traps in the local optima, eliminate premature convergence, effectively explore the search space and effectively use weaknesses. In addition, the improved algorithm provides the best or near-optimal solution, as well as the best compromise between research and the quality of algorithm use.

Several algorithms are described, which combine the best features of different algorithms to develop a new algorithm. Hafez et al. proposed the MAKHA algorithm, in which the jumping process is taken over from the monkey algorithm, and the evolution operator (mutation, crossover) of the krillherd algorithm is used to quickly find the optimal solution. The MAKHA algorithm was tested with the KNN classifier on 18 ICU data sets and achieved the classification accuracy.

One of the most popular and promising algorithm in the physics-based category is Simulated annealing (SA). Therefore, in order to improve the whale optimization algorithm, Mafarj and Mirjalili added SA to the WOA. They accelerate the use of WOA by improving the best solution after each solution. The performance of the hybrid WOASA algorithm was tested on 18 data sets using the KNN classifier.

The location update quality of the Crow search algorithm is used in the gray wolf optimizer, and a good balance is found between exploration and exploitation (Arora). They mixed an algorithm similar to GWOCSA, which was applied to the 21 known records in the UCI repository. Space uses transfer function S. The accuracy of the considered ANN classifier is compared with other meta-heuristic algorithms of the latest generation.

In order to get rid of the local optimum in the sine and cosine algorithm, Abd Elaziz proposed a hybrid algorithm using differential evolution algorithm for local search methods. An improved version of the sine-cosine algorithm was tested on eight ICU data sets and showed the best results in terms of performance and statistical analysis.

Tawhid and Dsuza developed a hybrid algorithm that uses the bat algorithm and an improved version of the PSO algorithm to solve the feature selection problem in binary space. In order to convert the position of the bat in the binary space, the V-shaped transfer function is used, and the S-form transfer function is used in the same way. The transfer function is used to obtain the binary position of the particle in the PSO. Get the algorithm. The hybrid HBBEPSO algorithm combines a good insight into the bat algorithm and the convergence characteristics of the PSO algorithm, and achieves the best performance from 20 standard data sets. Compared with all algorithms, the accuracy is high. Binary black hole algorithm and PSO algorithm are used to select the pre-processing method of genes for cancer classification, in order to more effectively improve the research and utilization rate. Various classifiers used to evaluate performance accuracy and obtain results from two standard microarrays and three clinical data sets.

Neggaz et al. proposed a hybrid algorithm that uses a sine-cosine algorithm to control a group of Salps to solve the feature selection problem. Baliarsingh built a social engineering optimizer in EPO to improve the performance of EPO. SVM classifier modified by memetic algorithm. And apply it together with the hybrid method proposed for medical records. The proposed hybrid algorithm is compared with other known meta-heuristic algorithms and is superior to other algorithms.

Another hybrid EPO method proposes a face recognition cultural algorithm. This method improves the ability of existing methods and is applied to face recognition together with the SVM classifier. It gives better results in terms of convergence and reliability. The sine cosine algorithm improves the research stage and avoids premature convergence.

In order to obtain the optimal gene from gene expression data, Shukla combines TLB with the SA algorithm. The SA algorithm is used to improve the solution quality of the TLBO algorithm and help find suitable cancer detection genes. A function to convert variables into binary variables is proposed. The SVM classifier was used to evaluate the classification accuracy and tested on ten sets of microarray data sets.

Various hybridizations of different meta-heuristic algorithms have been developed to solve different applications of feature selection problems. Jaya algorithm is used for genetic selection of forest optimization algorithm. The use of enhanced JA is used to optimize the two parameters of the forest optimization algorithm. Gray wolf optimizer and Grasshopper optimization algorithm are used to select text function, PSO algorithm and gravity search algorithm are used in industrial foam injection process. Use random fractal search algorithm and optimization algorithm of locust and cat colony to select features.

Below have been shared tables of modification and hybridization of both WOA and GWO algorithms:

Table 1 WOA modifications

Modification Name	Purpose	Conclusion
WOA in Neural Networks	WOA is used as an optimizer to control weight and biases in neural networks	Results presented that neural network by using WOA performs better compared to the Backpropagation algorithm
Chaotic WOA	Chaos was used to control the status of WOA and to improve the performance of convergence speed, and achieve a better result	Ten maps were tested in order to develop a chaotic set. CWOA improved the efficiency of WOA and balances between exploration and exploitation by using 0.7 as an initial point
Memetic WOA	Avoiding local optima is drawback of WOA. Therefore, MWOA was proposed in order to prevent WOA from this problem.	MWOA added a chaotşc search embedded inside the exploration phase and creates stability between exploration and exploitation
ILWOA	The decreasing cloud physical machine number qas the aim of improving ILWOA due to the available bandwidth.	ILWOA was tested on 25 mathematical functions and then the resault compared to WOA, which showed that ILWOA improved WOA performance
IWOA	The control parameter a is linear, so, it cannot work well with nonlinear problems inside the search process. IWOA used some nonlinear strategies to overcome this problem.	The result showed that IWOA performed well compared to standard WOA in convergence speed

Table 2 WOA Hybridizations

WOA Hybridization with	Purpose	Conclusion
BAT	Improving the exloration of WOA and obtaining a better solution in the exploitation phase was the aim of WOA-BAT	The WOA-BAT improved the quality of results against standard WOA and other algorithms
Artificial Neural Network based on WOA	Using WOA to overcome the balancing difficulties related to parameter settings	Results of the neural network based on WOA showed better performance, which is 9.9 % accuracy
PSO	The aim of PSO-WOA was to obtain better results for solving numerical functions that are global	PSO embedded inside the hunting phase and the result was more efficient compared to the standard WOA
BS (Brain Storm)	Privacy is a big challenge in cloud computing, so the secret key of data was identified by BS-WOA	Results showed that BS-WOA obtained better security by protecting the confidentiality and effectiveness of data in the cloud
CBO (colliding bodies optimization)	The aim od WOA-CBO was to improve the accuracy result, reliability and convergence speed	WOA-CBO compared with the standard WOA and results showed that WOA-CBO performed better than WOA
MFO	Avoiding time consuming for determining the best optimal thresholding in multi-threshold was the aim of WOA-MFO	Compared with five algorithms, WOA-MFO showed a better result in terms of speed, the best fitness value and the ANOVA test
LS (Local Search)	Reducing computational cost and avoiding local optima	The best result could be achieved quickly by using various techniques like swamp mutation, local search strategy and insert-reversed block

Table 3 GWO Modifications

Modifications	Purpose	Conclusion
Modified GWO	MGWO was proposed to tune recurrent neural network parameters, which then used for classifying students performance	As a result, MGWO could find the best solution than other competitive models. MGWO has a greater impact to improve the result of a recurrent neural network
Chaotic GWO	Increasing the convergence speed was the purpose of this modification by adding different chaotic methods	10 chaotic maps were used and the best was chosen in order to use it with GWO. The results showed that CGWO improved GWO
Binary GWO	There are different large-scale problems. Unite commitment problem was one of these problems that could be solved by BGWO	It was compared to the standard GWO and a variety of binary algorithms. The results showed that BGWO outperformed well
Intelligent GWO	The aim of IGWO was at solving different problems (predict the future energy price) in companies, which sale power in the energy market.	IGWO was tested on 22 benchöark functions. IGWO was compared with GWO, Oppositipnal GWO and PSO. Results showed IGWO superior to other algorithms
GWO	The aim of this work was to improve performance of the human recognition system	Results are compared to GA and Firefly Algorithm. The GWO outperformed well compared to GA and FA
Power GWO	GWO was used to solve complex optimization problems based on power local optimization approach, which was essential for clustering	PGWO was tested on 7 benchöark functions and tested on 9 data sets for clustering. Results showed that PGWO performed well against the most recent algorithms

Table 4 GWO Hybridizations

GWO Hybridizations with	Purpose	Conclusion
Dragonfly (DA)	The renewable energy system has some problems, such as voltage deviation, power loss and decreasing fuel cost, solution of which is the aim of this approach	The result showed that it was faster and improved its performance when the IEEE 30 bus system was used to test
Recurrent Neural Network	The learning experince and forecasting outcome of the students results was the aim of this hybridization	Results proved that the hybridized system improved the forecasting task in terms of accuracy compared to other models
Long Short Term Mermory (LSTM)	The recurrent neural network has some drawbacks related to accuracy, convergence speed. GWO was used to train the LSTM recurrent neural network	Simulation results presented that GWO can improve the performance of the recurrent neural networks by training the LSTM recurrent neural networks
Fireworks Algorithm (FWA)	The aim of this hybridization was to combine the 2 most efficient algorithms	The FWA-GWO was tested on 22 benchmark functions and been compared to FWA and GWO. The results showed that FWA-GWO outperformed other two standard algorithms
Flower Pollination Algorithm (FPA)	Hybridizing both algorithms to have a better solution in solving real-world applications was the aim of this hybridization	The hybridized approach was verified on 6 benchmark functions and then compared against PSO, FPA and GWO. GWO-FPA showed superiority in its performance
Sine Cosine Algorithm (SCA)	To improve the quality solution of GWO, GWO was hybridized with SCA	The results were compared with standard GWO, SCA, WOA, ALO and PSO. The GWO-SCA performed well in solving test functions and real-world problems
GA	Solving the economic dispatch problems was the aim of this approach	GWO was hybridized with a crossover and mutation mechanism for improving the performance. The results showed equality in some cases and better results in others

CHAPTER III

RESEARCH METHODOLOGY

This section describes the proposed method of the thesis. The main goal of this study is to propose a hybrid metaheuristic algorithm. Besides, image features are one of the big datasets in the science and engineering. The optimal image feature selection is also important, since the optimal feature selection is one of the NP-hard (nondeterministic polynomial) problems, the proposed hybrid method is applied on an image data set. The proposed hybrid algorithm is made by using the SSA, GWO and WOA algorithms. These algorithms are the well-known algorithms also the performance of each of them is significant. In this way, the mechanism of each algorithm is used on the hybrid algorithm. The hybrid algorithm follows the main structure of these algorithm since each algorithm has specific nature behavior.

3.1 Description of Proposed Method

The metaheuristic algorithm tries to find the optimal solution. The search agents in the metaheuristic algorithms select the optimal solution from the search space. They follow a mechanism to update the search space. One of the main problems in the heuristic algorithm is trapping on the local optima. The metaheuristic algorithm avoids local optimum trapping by the two phases; exploration and exploitation. In the exploration phase, the search agents search on the environment (area) to find an approximate solution and the optimal solutions which are found on the exploration phase are selected on the exploitation phase.

The hybrid metaheuristic algorithm mainly used two algorithms, newly the hybrid metaheuristic algorithms used three different metaheuristic algorithm. The proposed algorithm combines different nature behaviors of each algorithm to find optimal solution. As mentioned before, the main goal of the metaheuristic algorithm is trapping on local optimum avoidance, also the hybrid algorithms try to improve performance on trapping on local optimum avoidance. The metaheuristic algorithm not guarantee to find the best solution while they try to find optimal solution as well as the hybrid metaheuristic algorithms. Each algorithm is described in the previous section in depth.

In the SSA algorithm the leader search agent has the main role in the algorithm, in this way the other search agents follow the leader behavior and position. If the leader search agent chose the good solution the search agents remaining on the swarm select this solution to follow, on the other hand the bad selection of the leader effects the whole swarm. The GWO algorithm is the well-known algorithm but has some disadvantages such as; low solving precision, slow convergence, and bad local searching ability. The WOA algorithm suffers from improving approximate solutions during in each iteration in this way the WOA in the exploitation phase has not respectable performance. The proposed hybrid algorithm uses the advantages of the SSA, WOA, and GWO algorithms. The disadvantage of these algorithms are also improved in the hybrid algorithm. The main goal of hybrid algorithm is to avoid trapping on the local optimum, besides improve the exploration and exploitation phases, which is the one the issues that faced in the metaheuristic.

The hybrid algorithm consists of the three metaheuristic algorithm, SSA, WOA and GWO algorithms. As known in the metaheuristic algorithm there is a search space that initialized from random solutions. The search agents include salp swarm, whales, and grey wolves try to find optimal solution from the search space. The fitness value for each of the search agents is calculated based on the objective. There are three level fitness value, first, second and third based on the objective function is maximizing or minimizing the problem the fitness values are ordered.

Initially, the *i* search agents (i=1, 2, …, n) are divided two groups. In the first group the i^{th} search agent in the search space updates the current position based on the SSA algorithm structure (Eq.22). The remaining search agents between the second and $i/2^{th}$ update the position based on the GWO algorithm. In the GWO algorithm the position update is followed by the alpha, beta, and delta wolves. In this way, the hybrid algorithm update the omega wolves based on the first three wolves. The second group of the search agents in the hybrid algorithm follows the WOA algorithm structure. As mentioned in the previous section on thesis, the WOA algorithm consists of the two condition on position update. The random number between zero and one is selected. If the random number is smaller than 0.5 the search agent update, follow the eq. 30 or 31 on the other hand the search agents use the Eq. 32.

$$X_j^i = X_{Best_j} + c_1\left(\left(ub_j - lb_j\right)c_2 + lb_j\right) \quad (22)$$

$$\overrightarrow{D_\alpha} = |\overrightarrow{C_1}\,\overrightarrow{X_\alpha} - \vec{X}| \quad (23)$$

$$\overrightarrow{D_\beta} = |\overrightarrow{C_2}\,\overrightarrow{X_\beta} - \vec{X}| \quad (24)$$

$$\overrightarrow{D_\delta} = |\overrightarrow{C_3}\,\overrightarrow{X_\delta} - \vec{X}| \quad (25)$$

$$\overrightarrow{X_1} = |\overrightarrow{X_\alpha} - \overrightarrow{A_1}\,\overrightarrow{D_\alpha}| \quad (26)$$

$$\overrightarrow{X_2} = |\overrightarrow{X_\beta} - \overrightarrow{A_2}\,\overrightarrow{D_\beta}| \quad (27)$$

$$\overrightarrow{X_3} = |\overrightarrow{X_\delta} - \overrightarrow{A_3}\,\overrightarrow{D_\delta}| \quad (28)$$

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1}(t) + \overrightarrow{X_2}(t) + \overrightarrow{X_3}(t)}{3} \quad (29)$$

$$\vec{D} = |\vec{C}\,\overrightarrow{X^*}(t) - \vec{X}(t)| \quad (30)$$

$$\vec{X}(i+1) = \overrightarrow{X_{rand}} - \vec{A}\vec{D} \quad (31)$$

$$\vec{X}(i+1) = e^{bk}\cos(2\pi k)\,\overrightarrow{D^*} + \overrightarrow{X^*}(i) \quad (32)$$

As mentioned before, the image data set has many features. Some of these features maybe not important to save and use on the other studies. If the data set with many feature is used in the machine learning there are many time and cost to train the features to the machine. In this way, the reduced data set is important to use. Besides the reduced data set should be benefits the more effective features which has the significant role in the image retrieval. In this thesis, the hybrid algorithm is used to reduce the ionosphere data set. The reduced data set with more effective feature is evaluated by the two well-known classifier like SVM and KNN.

SVM classifier is a kernel-based supervised learning algorithm that can divide data into two or more classes. SVM is specially developed for binary classification. In the training phase, the SVM creates a model that displays the decision limits of each category and displays the hyperplane that separates the different categories. Increasing the distance between classes by increasing the hyperplane field helps to improve the classification accuracy.

The KNN algorithm is an instance-based learning method, which is used to classify objects according to their closest training instance in the feature space. Objects are classified by the majority decision of their neighbors, that is, the object is assigned to the category with the highest frequency among its nearest neighbors, where k is a positive integer. In the KNN algorithm, the classification of the new vector test attribute is determined by the classification of its "most complex" neighbor.

Feature selection can be modeled as a general optimization problem, in which two goals must be achieved; the minimum number of selection functions and the maximum classification accuracy. The optimal solution is to select a solution with few features and high classification accuracy. The KNN classification accuracy is used as a fitness function to evaluate the performance of all search agents.

In this thesis, the proposed hybrid algorithm has been used as a wrapper feature selection mode by applying the two of above mentioned classifiers as a fitness function. The proposed algorithm can be used with any other classifiers too. SVM and KNN are used one by one as a classification algorithm to evaluate the quality of the selected features. It can be seen from Figure 11 that the proposed model consists of two main stages; feature selection and classification. The model first takes the original data set as input, then the proposed hybrid algorithm tries to select valuable features, and then uses the obtained features to feed once the KNN and then the SVM classifier. The results have been compared and shared in the next section.

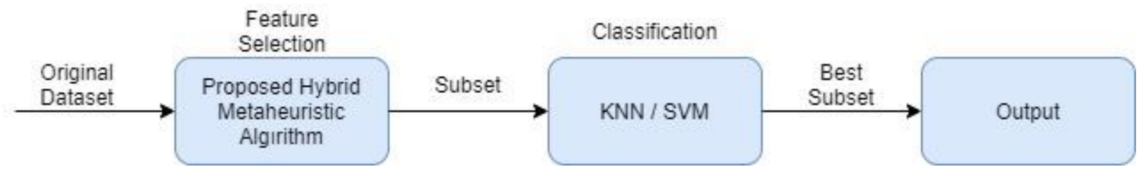

Figure 11 Stages of Feature Selection

3.2 Pseudocode of Proposed Hybrid Algorithm

```
Initialize Population
while iter <= maxIter
        Check if any agents go beyond the search space
        Calculate the Fitness values for each search agent
        Calculate Alpha, Beta and Delta
        Update c1 param
        if1 i<=N/2
            if2 i == 1
                Update the Leader using SSA (Eq. 22)
            else if2 i!=1
                Update the Followers of the Leader group using GWO (Eq. 23-29)
            end if2
        else if1 i> N/2
            Calculate the Best Fitness value of search agents using WOA
            if3 P<0.5
                if4 A>=1
                    Update the position of current search agent by Eq. 31
                else if4 A<1
                    Update the position of current search agent by Eq. 30
                end if4
            else if 3 P>= 0.5
                Update the position of current search agent by Eq. 32
            end if3
        end if1
end while
return results
```

Figure 12 Pseudocode of Proposed Hybrid Algorithm

3.3 Flowchart of Proposed Hybrid Algorithm

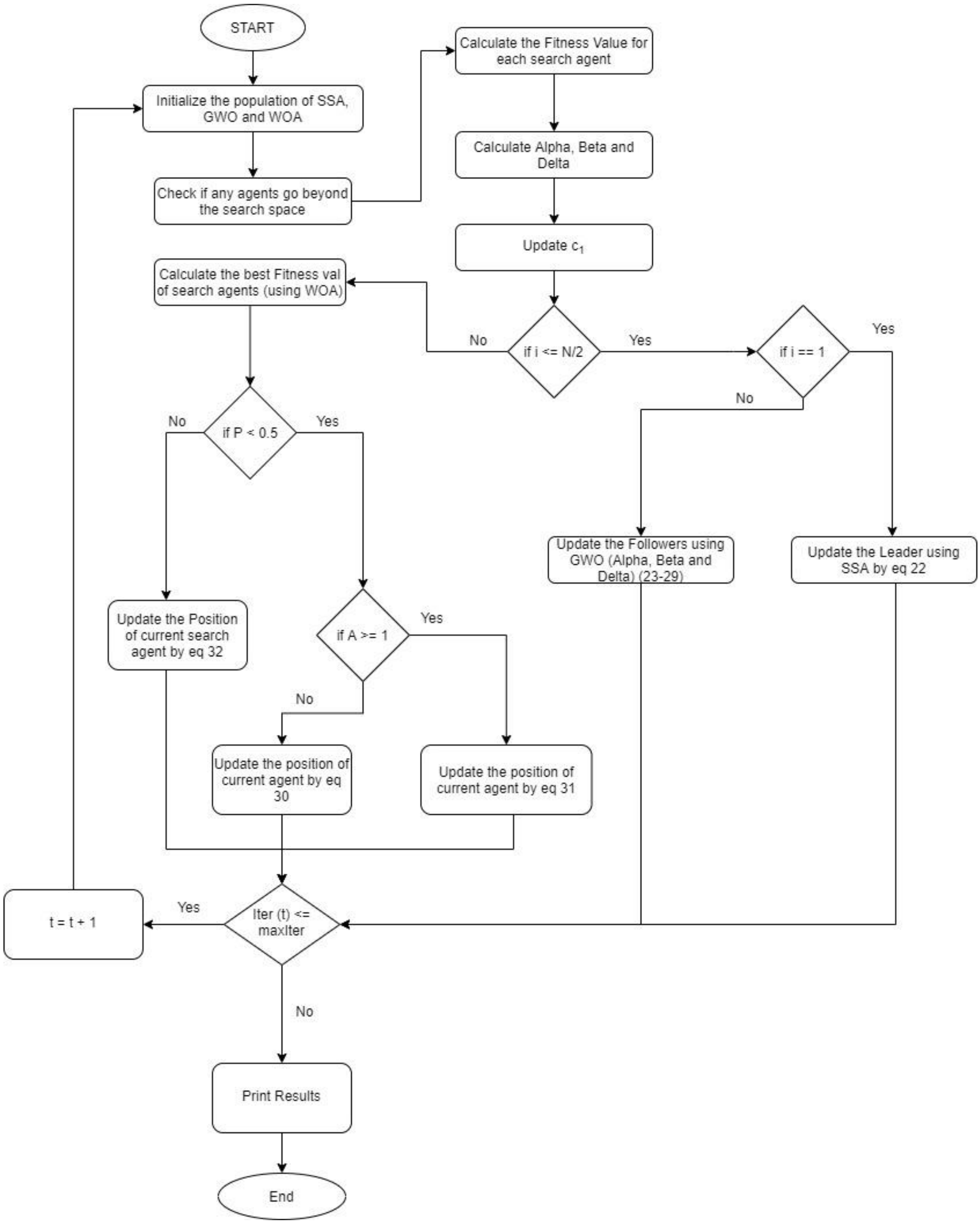

Figure 13 Flowchart of Proposed Hybrid Algorithm

CHAPTER IV

FINDINGS AND DISCUSSION

As mentioned before, the optimal features in data set is obtained. In this section we want to discuss and evaluate about reduced data set. The importance of this chapter is due to the simulation results performed in Matlab 2016 software. This software has a high ability in programming ultra-innovative algorithms. It will also be seen in this chapter that we have reached the optimal solutions by implementing an optimization algorithm based on proposed hybrid algorithm.

The ionospheric data set from the UCI machine learning database repository is used to evaluate the impact of feature selection methods on data classification. Ionospheric Dataset represents the data obtained from the classification of the ionospheric radar reflection signal. There are two classes of "good" and "bad", and each instance contains 34 attributes. As mentioned above, two shell-based methods have been used to create the best subset of functions as classifiers.

The main data set has 34 features and 351 records. Besides, the 35th column shows the result of the feature row being good or bad. With the help of four optimization algorithms, we have been able to make this database smaller. The main purpose of this dissertation is to reduce the specificity of the main data set to get solution close to the main database, for example, the time complexity and cost in calculating of 34 x 351 database is much more than calculating a database in Dimensions are 6 x 351. In the table 5, an example is presented.

Table 5 Results of feature selection for all optimizers

Algorithm Name	**No. of features**	**No. of selected**	**Feature indices**
Proposed algorithm	34	7	5,7,8,18,24,27,34
WOA	34	7	1,7,8,11,18,23,34

SSA	34	6	5,6,9,13,23,31
GWO	34	7	1,6,8,12,24,27,34

In this thesis, the proposed method based on a hybrid metaheuristic algorithm is applied on ionosphere data set to reduce image features. In this way, the data set is reduced to an optimal dataset with some features that has main role on the image classification. This study follows two data classifier algorithm named as SVM and KNN. The obtained result is the accuracy rate of this two algorithm. The hybrid metaheuristic algorithm is combined with SSA, GWO and WOA algorithms. To evaluate the obtained results each classifier algorithm runs 30 times. The metaheuristic algorithm simulation parameters are presented in the table 6.

Table 6 The Simulation Parameters

Algorithm	Parameter	value
GWO	a	[2,0]
WOA	a	[2,0]
	A	[2,0]
	C	2.rand(0,1)
	l	[-1,1]
	b	1
SSA	C_2 , C_3	[0,1]

We used random selection to distribute the records because the learning functions must first acquire knowledge of the properties and the results of the properties and then extract new information from the knowledge gained. As mentioned, we have 351 feature records, so a number of these records must be selected as knowledge. That we can determine the percentage that the result of learning is directly related to this percentage, that is, the more records learned, the closer we are to the optimal answer. Therefore, we have done the learning algorithms once the learning rates of 0.3, 0.5 and 0.8 to have a comparison in terms of learning rates.

Finding the optimal answer is calculated by calculating the area under the curve, which is a numerical value between zero and one. The closer the value is to one, the more predictable the learning outcome is. Now, after executing the learning algorithms, Table 7, 8, 9 presents the obtained results. The accuracy of each algorithm is calculated by the sum of TP (True positive) and TN (True negative) divided by the TP, TN, FP (False positive), FN (False negative) that given by the Eq. 33:

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (33)$$

Table 7 KNN and SVM classifier with 0.3 train rate. (30 runs)

Algorithm	**Classifier**	**Average**	**Standard division**
Proposed Method	KNN	**0.793**	4.1E-02
	SVM	**0.781**	3.32E-02
SSA	KNN	0.763	3.2E-02
	SVM	0.775	3.1E-02
WOA	KNN	0.761	3.33E-02
	SVM	0.769	6.31E-02
GWO	KNN	0.788	2.82E-02
	SVM	0.775	3.65E-02

While the train rate is 0.5, the proposed approach is superior in KNN classifier, while in the SVM the GWO seems to be superior than the other algorithms, including the proposed approach too.

Table 8 KNN and SVM classifier with 0.5 train rate. (30 runs)

Algorithm	**Classifier**	**Average**	**Standard division**
Proposed Method	KNN	**0.816**	3.33E-02
	SVM	0.772	1.02E-02
SSA	KNN	0.810	2.27E-02
	SVM	0.806	4.14E-02
WOA	KNN	0.793	3.15E-02
	SVM	0.796	4.91E-02
GWO	KNN	0.811	2.98E-02
	SVM	**0.823**	1.51E-02

In the another part of the discussion on the results of this thesis, we increased the random selection variable to 0.8 which means increasing the number of learing records. In other words, if we have 400 records in the database and considering 0.8 as the rate Learning means that 320 records from this database are considered as training. The remaining 80 records will be considered as experiments. In this section, due to the randomness of the selection of two databases on four learning algorithms, it is repeated 30 times and the results consists of an average of 30 repetitions. At the table 9 the results for the 0.8 train rate are shown, and as we can see, the proposed approach is superior in both of the classifiers with the highest accuracy compared to other algorithm.

Table 9 KNN and SVM classifier with 0.8 train rate. (30 runs)

Algorithm	**Classifier**	**Average**	**Standard division**
Proposed Method	KNN	**0.843**	1.25E-02
	SVM	**0.853**	2.12E-02
SSA	KNN	0.821	7.95E-02
	SVM	0.833	4.2E-02
WOA	KNN	0.816	4.31E-02
	SVM	0.824	1.26E-02
GWO	KNN	0.833	4.36E-02

	SVM	0.837	3.24E-02

Here, to clarify the evaluation the obtained results are shown in the figures 14,15,16. As shown in the figure, the proposed method has good performance in comparison other metahehuritis algorithms. The obtained results also clarify the KNN and SVM algorithm get optimum results.

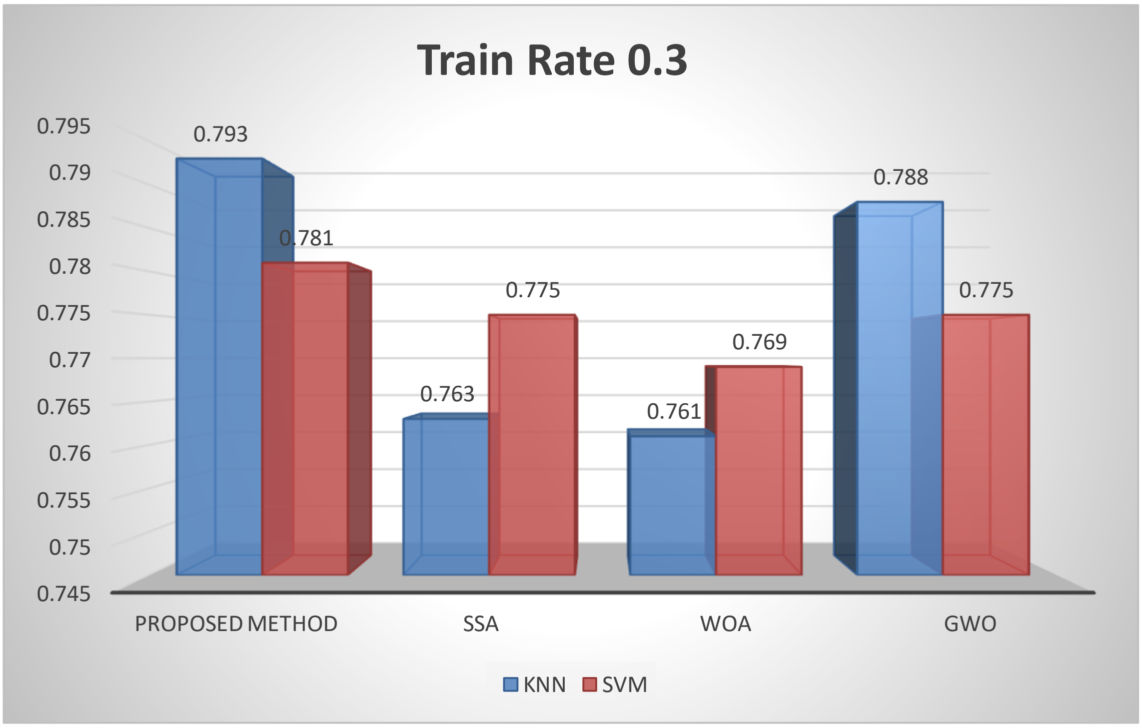

Figure 14 Comparison by 0.3 Train Rate

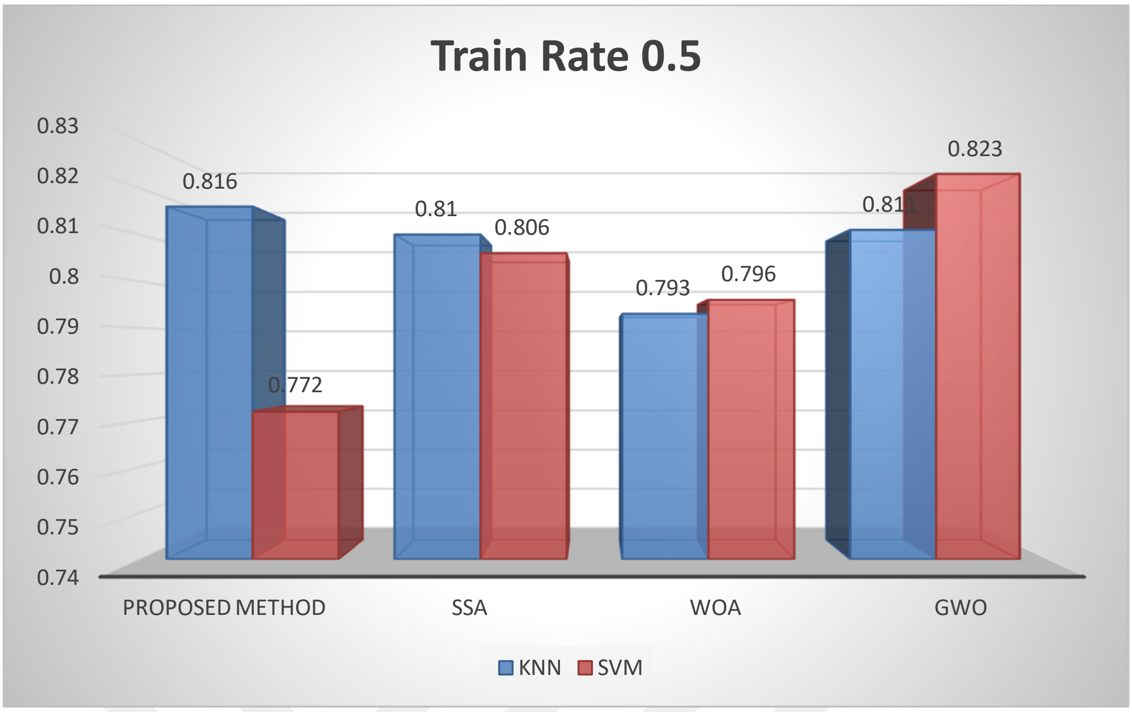

Figure 152 Comparison by 0.5 Train Rate

Figure 16 Comparison by 0.8 Train Rate

CHAPTER V

CONCLUSION

The process of selecting a subset of features, which can maximize the performance of the learning algorithm and reduce the dimensionality of the problem space is known as feature selection. In this study, a new and efficient approach for image feature reduction have been proposed based on a new hybrid metaheuristic algorithm. This approach aimed to optimally select the best feature set and obtain comparable or better classification accuracy from using all features of images. The goal of this study is reached by the new hybrid metaheuristic algorithm which also was proposed in this study. It consisted of hybridizing the SSA, GWO and WOA nature inspired algorithms to optimally select the best feature set of the dataset. The performance of the proposed algorithm has been evaluated and compared with the basic metaheuristic algorithms of SSA, GWO and WOA. The proposed approach has outperformed the other algorithms in classification accuracy, being the one that has selected the best minimal number of features. The results showed that our method can reduce the image datasets by finding optimal combination of features with maximum classification accuracy. The main role of these results has been played by the new hybrid metaheuristic algorithm which based on the results demonstrated that achieves a better performance than other algorithms.

REFERENCES

Mirjalili, S., Gandomi, A. H., Mirjalili, S. Z., Saremi, S., Faris, H., & Mirjalili, S. M. (2017). Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Advances in Engineering Software*, *114*, pp. 163-191.

Ling, Y., Zhou, Y., & Luo, Q. (2017). Lévy flight trajectory-based whale optimization algorithm for global optimization. *IEEE access*, *5*, pp. 6168-6186.

Seyyedabbasi, A., & Kiani, F. (2019). I-GWO and Ex-GWO: improved algorithms of the grey wolf optimizer to solve global optimization problems. *Engineering with Computers*, pp. 1-24.

Seyyedabbasi, A., Aliyev, R., Kiani, F., Gulle, M. U., Basyildiz, H., & Shah, M. A. (2021). Hybrid algorithms based on combining reinforcement learning and metaheuristic methods to solve global optimization problems. *Knowledge-Based Systems*, 223, 107044.

Haklı, H., & Uğuz, H. (2014). A novel particle swarm optimization algorithm with Levy flight. *Applied Soft Computing*, *23*, pp. 333-345.

Mirjalili, S. (2016). SCA: a sine cosine algorithm for solving optimization problems. *Knowledge-based systems*, *96*, pp. 120-133.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, *1*(1), pp. 67-82.

Eberhart, R., & Kennedy, J. (1995). Particle swarm optimization. In *Proceedings of the IEEE international conference on neural networks* (Vol. 4, pp. 1942-1948). Citeseer.

Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, *95*, pp. 51-67.

Yang, X. S. (2010). Firefly algorithm, stochastic test functions and design optimisation. *International journal of bio-inspired computation*, *2*(2), pp. 78-84.

Sayed, G. I., Khoriba, G., & Haggag, M. H. (2018). A novel chaotic salp swarm algorithm for global optimization and feature selection. *Applied Intelligence*, *48*(10), pp. 3462-3481.

Abbassi, R., Abbassi, A., Heidari, A. A., & Mirjalili, S. (2019). An efficient salp swarm-inspired algorithm for parameters identification of photovoltaic cell models. *Energy conversion and management*, *179*, pp. 362-372.

Ibrahim, R. A., Ewees, A. A., Oliva, D., Abd Elaziz, M., & Lu, S. (2019). Improved salp swarm algorithm based on particle swarm optimization for feature selection. *Journal of Ambient Intelligence and Humanized Computing*, *10*(8), pp. 3155-3169.

Ahmed, S., Mafarja, M., Faris, H., & Aljarah, I. (2018, March). Feature selection using salp swarm algorithm with chaos. In *Proceedings of the 2nd International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence* (pp. 65-69).

Neggaz, N., Ewees, A. A., Abd Elaziz, M., & Mafarja, M. (2020). Boosting salp swarm algorithm by sine cosine algorithm and disrupt operator for feature selection. *Expert Systems with Applications*, *145*, 113103.

Tubishat, M., Idris, N., Shuib, L., Abushariah, M. A., & Mirjalili, S. (2020). Improved Salp Swarm Algorithm based on opposition based learning and novel local search algorithm for feature selection. *Expert Systems with Applications*, *145*, 113122.

Panda, N., & Majhi, S. K. (2020). Improved salp swarm algorithm with space transformation search for training neural network. *Arabian Journal for Science and Engineering*, *45*(4), 2743-2761.

Abd Elaziz, M., Li, L., Jayasena, K. N., & Xiong, S. (2020). Multiobjective big data optimization based on a hybrid salp swarm algorithm and differential evolution. *Applied Mathematical Modelling*, *80*, 929-943.

Faris, H., Mafarja, M. M., Heidari, A. A., Aljarah, I., Ala'M, A. Z., Mirjalili, S., & Fujita, H. (2018). An efficient binary salp swarm algorithm with crossover scheme for feature selection problems. *Knowledge-Based Systems*, *154*, 43-67.

Mirjalili, S., Hashim, S. Z. M., & Sardroudi, H. M. (2012). Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm. *Applied Mathematics and Computation*, *218*(22), 11125-11137.

Mirjalili, S., & Lewis, A. (2013). S-shaped versus V-shaped transfer functions for binary particle swarm optimization. *Swarm and Evolutionary Computation*, *9*, 1-14.

Liang, J. J., Qin, A. K., Suganthan, P. N., & Baskar, S. (2006). Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. *IEEE transactions on evolutionary computation*, *10*(3), 281-295.

Yang, X. S., & He, X. (2016). Nature-inspired optimization algorithms in engineering: overview and applications. *Nature-inspired computation in engineering*, 1-20.

Michalewicz, Z., & Fogel, D. B. (2004). Constraint-Handling Techniques. In *How to Solve It: Modern Heuristics* (pp. 231-270). Springer, Berlin, Heidelberg.

Wang, G., & Guo, L. (2013). A novel hybrid bat algorithm with harmony search for global numerical optimization. *Journal of Applied Mathematics*, *2013*.

De Giovanni, L., & Pezzella, F. (2010). An improved genetic algorithm for the distributed and flexible job-shop scheduling problem. *European journal of operational research*, *200*(2), 395-408.

Salman, A., Ahmad, I., & Al-Madani, S. (2002). Particle swarm optimization for task assignment problem. *Microprocessors and Microsystems*, *26*(8), 363-371.

Tate, D. M., & Smith, A. E. (1995). A genetic approach to the quadratic assignment problem. *Computers & Operations Research*, *22*(1), 73-83.

Srinivas, M., & Patnaik, L. M. (1994). Genetic algorithms: A survey. *computer*, *27*(6), pp. 17-26.

Eberhart, R., & Kennedy, J. (1995, October). A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science* (pp. 39-43). IEEE.

Teodorović, D. (2009). Bee colony optimization (BCO). In *Innovations in swarm intelligence* (pp. 39-60). Springer, Berlin, Heidelberg.

Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in engineering software*, *69*, 46-61.

Mohammed, H. M., Umar, S. U., & Rashid, T. A. (2019). A systematic and meta-analysis survey of whale optimization algorithm. *Computational intelligence and neuroscience*, *2019*.

Emary, E., Zawbaa, H. M., Grosan, C., & Hassenian, A. E. (2015). Feature subset selection approach by gray-wolf optimization. In *Afro-European conference for industrial advancement* (pp. 1-13). Springer, Cham.

Sharawi, M., Zawbaa, H. M., & Emary, E. (2017, February). Feature selection approach based on whale optimization algorithm. In *2017 Ninth international conference on advanced computational intelligence (ICACI)* (pp. 163-168). IEEE.

Braga, P. L., Oliveira, A. L., & Meira, S. R. (2008, March). A GA-based feature selection and parameters optimization for support vector regression applied to software effort estimation. In *Proceedings of the 2008 ACM symposium on Applied computing* (pp. 1788-1792).

Huang, C. L., & Dun, J. F. (2008). A distributed PSO–SVM hybrid system with feature selection and parameter optimization. *Applied soft computing*, *8*(4), 1381-1391.

Zawbaa, H. M., Hassanien, A. E., Emary, E., Yamany, W., & Parv, B. (2015, December). Hybrid flower pollination algorithm with rough sets for feature selection. In *2015 11th International Computer Engineering Conference (ICENCO)* (pp. 278-283). IEEE.

Hafez, A. I., Hassanien, A. E., Zawbaa, H. M., & Emary, E. (2015, December). Hybrid monkey algorithm with krill herd algorithm optimization for feature selection. In *2015 11th international computer engineering conference (ICENCO)* (pp. 273-277). IEEE.

Agrawal, P., Abutarboush, H. F., Ganesh, T., & Mohamed, A. W. (2021). Metaheuristic Algorithms on Feature Selection: A Survey of One Decade of Research (2009-2019). *IEEE Access*, *9*, 26766-26791.

CURRICULUM VITAE (CV)

Personal Information

Name Surname:	Armir Kaçabeti
Nationality:	Albanian

Education Information

09/2012 – 06/2016	Bachelor – Computer Engineering Istanbul Sabahattin Zaim University
09/2017 – 06/2021	Master of Science – Computer Science and Engineering Istanbul Sabahattin Zaim University

Work Experience

06/2016 – 01/2018	Software Developer Emiturk Elektronik Sanayi ve Ticaret Ltd. Şti.
01/2018 – Present	Software Developer DATAPROM – Karan Bilişim Yazılım Güvenlik Sistemleri

Research Area

Machine Learning, Metaheuristic Algorithms, Cyber Security