IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 29, 2020, accepted June 6, 2020, date of publication June 11, 2020, date of current version June 24, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.3001685*

# Dual Kidney-Inspired Algorithm for Water Quality Prediction and Cancer Detection

**SALWANI ABDULLAH[1] AND NAJMEH SADAT JADDI[1,2]**

[1]Center for Artificial Intelligence Technology, Data Mining and Optimization Research Group (DMO), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[2]Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran 14117-13116, Iran

Corresponding author: Najmeh Sadat Jaddi (najmehjaddi@gmail.com)

This work was supported in part by the Ministry of Education, Malaysia, under Grant FRGS/1/2019/ICT02/ UKM/01/1, and in part by the Universiti Kebangsaan Malaysia under Grant DIP-2016-024.

**ABSTRACT** The kidney-inspired algorithm (KA) was presented in a recent research paper as a metaheuristic search algorithm. The KA imitates the physiological process of the kidneys in the human body. The second kidney in the human body filters all the solutes if the other kidney fails. If the second kidney also gets damaged, dialysis can be performed as a treatment. The failure of a kidney is proved by the Glomerular Filtration Rate (GFR) calculation. In this paper, a Dual-population Kidney-inspired Algorithm (Dual-KA), which contributes to the enhancement of a diversity of solutions and, subsequently, better exploration of the search space, is proposed, as a research objective, in the form of a novel simulation of cooperation between kidneys in human body. If GFR is greater than 60 in each iteration, the process in Dual-KA is continued as normal. If this number is between 15 and 60, some treatment is performed. Else, if the GFR is less than 15 and the current sub-population is the first sub-population with a GFR of less than 15, all the solutions in the current sub-population are sent to the other sub-population in the current iteration. Otherwise, a dialysis or transplant process is carried out. Dual-KA is tested on 12 known test functions, standard classification and time series datasets. Water quality prediction and cancer detection are two real-world applications of this algorithm. The mentioned algorithm is run on these problems and its high ability is proved.

**INDEX TERMS** Dual kidney-inspired algorithm, artificial neural network, water quality prediction, cancer detection.

## I. INTRODUCTION

Optimization methods, which are normally provided by metaheuristic algorithms as one of the intelligent system techniques, can optimize Artificial Neural Network (ANN) models. Many researchers have employed metaheuristic algorithms to train ANN models [1]–[9]. Finding the global optima solution is the most important goal of the optimization process. Interaction between exploration and exploitation is the main mechanism for finding the global optimum. However, a greater amount of exploration compared to exploitation aids the algorithm in finding the global optimum. In many research studies, it has been proved that a multi-population maintains the diversity of the solutions and enables movement towards the global optimum. In [10] a multi-population cooperative method for Particle Swarm Optimization (PSO) called CPSO-S was offered. This method was rooted in dividing the solution vector into smaller parts where each part is optimized separately. Another multi-population PSO-based method that engages two sub-populations searching separately in parallel with regular information passing between them was proposed in [11]. Other multi-population approaches applied to PSO can be found in the literature [12]–[17]. Some multi-population Genetic Algorithm (GA)-based approaches have also been proposed [15], [18]–[21]. Moreover, a ring-master-slave multi-population bat algorithm-based approach was performed in [4]. In addition, in [3] another multi-population method for cuckoo search was proposed which was used as a master-leader-slave strategy of multi-population. In this study, a master-slave strategy is guided using another sub-population called leader. Study of harmony search algorithm with multi-population that combines cooperative and competitive approaches was proposed in [22]. In line with these studies a unique strategy inspired from the physiological process of the two kidneys in human body is proposed with the aim of more exploration in basic KA [23].

The associate editor coordinating the review of this manuscript and approving it for publication was Fan-Hsun Tseng.

The method presented in this paper is a multi-population approach simulated from cooperation between the two kidneys in the human body, based on the KA proposed in [23]. The KA simulates the filtration mechanism of the human kidneys. The population of this algorithm contains solutes (solutions) filtered using the filtration rate (*fr*), which is computed based on the quality of solutions in the population. Good-quality solutions (with better quality compared to *fr*) are sent to Filtered Blood (FB) and poor-quality solutions (with poor quality compared to *fr*) are moved to Waste (W). Then, based on a number of conditions, reabsorption, secretion, and excretion are carried out. Finally, W and FB are combined and *fr* is updated. The KA has solved a number of optimization problems in previous research works [24]–[33].

In this paper, a novel simulation of how the two kidneys work together in the human body is presented to form a dual-population kidney-inspired algorithm, called Dual-KA. Dual-KA is proposed in order to achieve the research objective of enhancing the diversity of solutions and, subsequently, getting a better exploration of the search space for the KA algorithm. Cooperation between the kidneys in the human body results in a perfect blood filtration process if both kidneys are healthy [34]. In the biological system of kidneys in human body, checking the Glomerular Filtration Rate (GFR) is the best investigation to calculate the level of kidney function and, as a result, decide the stage of kidney disease. If GFR is greater than 60 in GFR medical test in the biological system of kidneys, the kidney function is normal. When one kidney gets damaged and the GFR reported in kidneys functionality test results is between 15 and 60, some medication and treatment is usually required. However, a significant problem arises when the GFR in kidneys functionality test is less than 15 as this denotes that one or both kidneys is failing to perform the filtration process. If one of the kidneys fails, the other kidney filters all the blood. If both kidneys fail, then dialysis or transplant is applied as a treatment. Dual-KA is a simulation of this cooperation between two kidneys. In Dual-KA, the GFR is checked in each iteration, each population. If the GFR is greater than 60, the process is continued as normal. If it is between 15 and 60, a kind of treatment is performed. However, if the GFR is less than 15 and the current sub-population is the first sub-population with a GFR of less than 15, in the current iteration, the entirety of solutions in the current sub-population are sent to the other sub-population. If the second kidney also fails, a dialysis or transplant process is carried out. By sending the solutions to the other sub-population, the sub-population with all the solutions can carry out the process with a new filtration rate. The contribution of these modifications is to enable the algorithm to better explore the search space. Cooperation between two kidneys in Dual-KA results in the enhancement of the diversity of solutions and, therefore, contributes in the improvement of the exploration of the algorithm which leads the algorithm towards superior results.

The remainder of this paper is structured as follows: Section 2 provides a brief description of the KA in literature. Section 3 presents the proposed method in detail, which involves simulating the cooperation between the two kidneys. Then, Section 4 presents a discussion of the results of the test functions, standard classification problems and also time series datasets. In addition, Section 4 presents the application of the proposed method in real-world water quality data and cancer detection problems. Section 5 concludes the study in this research.

## II. KIDNEY_INSPIRED ALGORITHM

An algorithm called KA that imitates the process of blood filtration and urine formation in the human body was presented in [23]. The algorithm works by mimicking four basic processes carried out by the kidneys: filtration, reabsorption, secretion, and excretion. In the first step of KA, a population of solutions is randomly created and initialized, the objective function of the solutions in the population is calculated and the best solution is selected. The movement of the solutions, which is simulated as new solution generation, is represented by Eq. (1):

$$Sol_{i+1} = Sol_i + \beta \times (Sol_{best} - Sol_i) \tag{1}$$

where a solution in the $i^{th}$ iteration is represented as $Sol_i$, best solution so far as $Sol_{best}$ and a random number $\beta$ in the range of [0, 1] is used. The filtration is processed in this algorithm by moving the solutions either to FB or W. This movement is performed with the aid of a threshold called the filtration rate (*fr*) and is calculated by Eq. (2):

$$fr = \alpha \times \frac{\sum_{i=1}^{p} f(x_i)}{p} \tag{2}$$

where $\alpha$ is constant value in the range of [0, 1], the objective function of solution $x$ in the $i^{th}$ iteration is represented by $f(x_i)$, and $p$ is the population size. The pseudocode of the entire process performed by KA is given in Fig. 1.

Reabsorption operation is performed when a potential new member of W can shift to FB if it meets the condition for inclusion in FB (passing the filtration operation). If not, the solution is excreted and removed from W (excretion operation) and a random number is inserted into W in replacement. This is with the aim of keeping the number of solutions fixed. Inserting a random number in the KA mechanism is inspired from the biological kidney system that solutes are regularly inserted into the kidneys system. Secretion operation from FB is performed when a new element of FB is not better than the worst solution in FB. However, if it is better than the worst solution, this solution is placed in FB and the worst solution is secreted and moved to W. This is secretion operation. Next, the best solution found so far is updated. As a final step, FB and W are merged to produce an integrated population for the next iteration. This procedure is continued until the termination criterion is satisfied.

```
set the population
evaluate the solute in the population
set the best solute, Sol_best
set filtration rate, fr, Eq. 2
set waste, W
set filtered blood, FB
set number of iterations, numOfIte
set ite ← 0
do while (ite<numOfIte)
        for all Sol_i
                generate new Sol_i, Eq. 1
                check the Sol_i using fr
                if Sol_i assigned to W
                        apply reabsorption and generate Sol_new, Eq. 1
                        if Sol_new cannot be a part of FB
                                remove Sol_i from W (excretion)
                                insert a random Sol into W to replace Sol_i
                        endif
                        Sol_new is reabsorbed
                else

                if it is better than the Sol_worst in FB
                                Sol_worst is secreted
                        else
                                Sol_i is secreted
                        endif
                endif
        endfor
        rank the Sols from FB and update the Sol_best
        merge W and FB
end while
return Sol_best
```

**FIGURE 1.** Pseudocode of kidney-inspired algorithm.

## III. DUAL KIDNEY_INSPIRED ALGORITHM

Although anybody can stay alive with only one kidney, the body functions is better overall when there are two. In the human body, the two kidneys are sent an equal amount of blood on which to perform the filtration process. The rate of blood filtration is the same on both sides in normal kidneys. Cooperation between kidneys results in a perfect blood filtration process if both kidneys are healthy (GFR $>$ 60). However, if there is a problem that adversely affects the activities of one of the kidneys, the other kidney will perform the actions of the ill-functioning kidney. When a kidney gets damaged and the GFR is between 15 and 60, some medication and treatment is required to support the poorly functioning kidney. However, a significant problem arises when the GFR is less than 15 in one or both kidneys and they fail to perform the filtration process properly. If one of the kidneys fails, the other kidney performs filtration on all the blood in the human body. However, if both kidneys fail, dialysis treatment or a transplant operation is required.

A simple simulation of this cooperation between the two kidneys is employed to develop Dual-KA. In other words, the difference between the proposed dual population and other multi-population algorithms is that in Dual-KA cooperation between populations is performed in a manner which is carried out between two kidneys in the human body. This is the novelty of the proposed method. To implement this novel cooperation in Dual-KA, a dual population of the basic KA is designed and developed, where each population is a representation of each kidney. In Dual-KA, the KA process is performed for each sub-population separately but in parallel in each iteration. After the basic process of KA has been performed for each sub-population, the GFR is checked. Based on different ranges of the GFR, a different treatment is applied. For minimization problems, the GFR is calculated based on Eq. 3, and for maximization problems, based on Eq. 4:

$$GFR_{\min} = 120 - \left(\frac{fr_{FB} * 100}{fr}\right) \tag{3}$$

$$GFR_{\max} = \frac{fr_{FB} * 12}{fr} \tag{4}$$

where, for both Eq. 3 and Eq. 4, the filtration rate for the current population is shown as $fr$, and the filtration rate for the population in FB is shown as $fr_{FB}$. According to the biological kidneys function, GFR level is always less than 120 in kidney functionality test results [34]. Based on this nature a simulation for GFR in Dual-KA for both minimization and maximization problems is performed. In this inspiration, Eq. 3 and Eq. 4 are given for minimization and maximization respectively. Since the data normalization in range of [0, 1] was performed to match artificial neural network, solutions in population are in the range [0, 1] too and filtration rate ($fr$) is in this range as a result. In minimization, since the quality of solutions in FB is lower than the quality of solutions

in population and it is required that GFR always stays less than 120, Eq. 3 is proposed, whilst in maximization, Eq. 4 is formulated when the quality of solutions in FB is greater than the quality of solutions in population, considering GFR is less than 120. Eq. 3 and Eq. 4 are the simplest form of the formulas (the reason of having 12 in Eq. 4), which can be proved as of Eq. 5 and Eq. 6:

$$\textit{In minimization } (fr_{\text{FB}} < fr)\&(0 < fr < 1)\&(0 < fr_{\text{FB}} < 1) \rightarrow (fr_{\text{FB}} \times 100)/fr \leq 120 \rightarrow GFR \leq 120 \quad (5)$$

$$\textit{In maximization } (fr_{\text{FB}} > fr)\&(0 < fr < 1)\&(0 < fr_{\text{FB}} < 1) \rightarrow GFR \leq 120 \quad (6)$$

In Dual-KA, if the GFR is greater than 60 the process is continued as normal (when the kidneys are working well). In this case, the algorithms are performing the filtration process well enough that there is no need to perform any treatment process. On the other hand, if the GFR is between 15 and 60, a kind of treatment needs to be performed because the quality of the solutions in FB is not good. Some treatment is needed to improve the quality of the solutions in FB and lead the algorithm towards a better searching process. In this case, Eq. 1 is applied to the solutions in FB in order to aid the algorithm in effectively exploring the search space. Figure 2 is a schematic representation of this process. In Fig. 2(a), kidney A has a GFR between 15 and 60 and kidney B works as normal, while Fig. 2(b) shows a situation in which kidney A works well and kidney B needs some treatment.

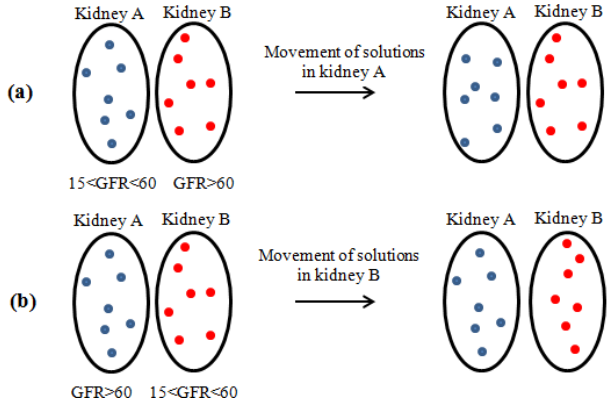


**FIGURE 2.** Treatment process when only one kidney has a GFR between 15 and 60: (a) kidney A with a GFR between 15 and 60, (b) kidney B with a GFR between 15 and 60.

When the GFR is less than 15 and the current sub-population is the first sub-population with a GFR of less than 15 in the current iteration, all the solutions in the current sub-population are sent to the other sub-population to proceed with the filtration. This is a simulation when one of the kidneys has failed in the human body and the other kidney takes on the entire filtration process on its own. Figure 3 shows this simple process schematically.

Fig. 3(a) illustrates the sending of the solutions in kidney A to kidney B and Fig. 3(b) shows vice versa. However, if the second sub-population also has a GFR of less than 15,

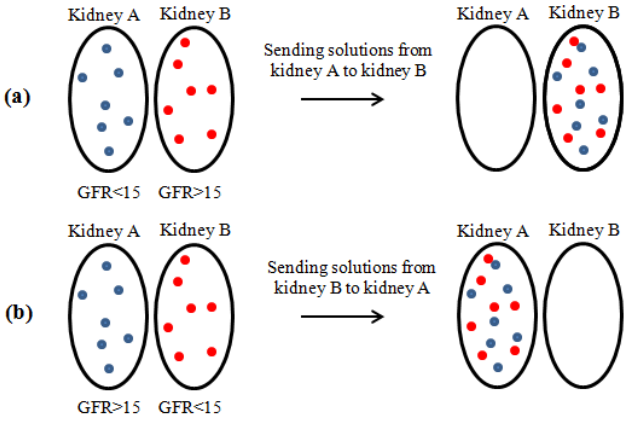


**FIGURE 3.** Treatment process when only one kidney has a GFR of less than 15: (a) kidney A with a GFR of less than 15, (b) kidney B with a GFR of less than 15.

either dialysis or a transplant needs to be performed based on the settings of the algorithm in the initial step. In this case, both sub-populations have poor solutions and no improvement is performed. Dialysis performs the KA process on the FB population, whereas a transplant restarts in one of the sub-populations with random solutions.

Fig. 4 illustrates the situation where both kidneys fail and either dialysis or a transplant is needed. In Fig. 4(a) kidney B is the second kidney with a GFR of less than 15 and in Fig. 4(b) kidney A is the second kidney with a GFR of less than 15. When the solutions are sent to the other sub-population, the sub-population with all the solutions will perform the KA process with a new filtration rate. The sub-population with a GFR of less than 15 has poor solutions. Therefore, when these solutions are sent to another sub-population containing higher-quality solutions, a mutation occurs to $fr$ and this may result in enhanced exploration being performed by the algorithm, enabling it to find a better solution in the search space.

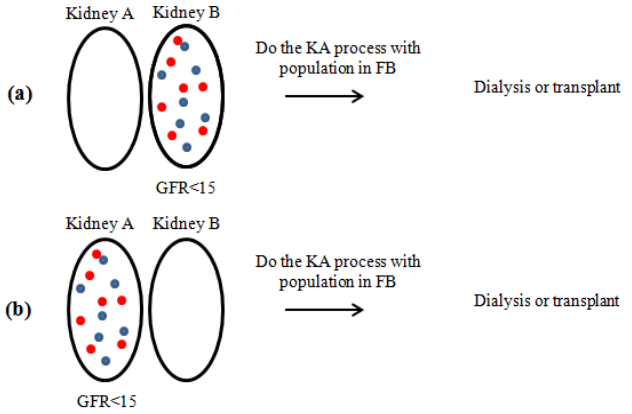


**FIGURE 4.** Treatment process when both kidneys have a GFR of less than 15: (a) kidney B as second kidney with a GFR of less than 15, (b) kidney A as second kidney with a GFR of less than 15.

The aim of applying the dialysis process is to improve the quality of the solutions in FB. The application of KA to the solutions in FB is a simulation of the filtration process using a

```
set the solutes in each kidney
evaluate the solutes in the kidneys
set the best solutes, Sol_best for each kidney
set filtration rate, fr, Eq. 2 for each kidney
set waste, W and filtered blood, FB for each kidney
set dialysis, Dia and transplant, Trans
set number of iterations, numOfIte
set ite ← 0
do while (ite<numOfIte)
    for two kidneys
        for all Sol_i
            generate new Sol_i, Eq. 1
            check the Sol_i using fr
            if Sol_i assigned to W
                apply reabsorption and generate Sol_new, Eq. 1
                if Sol_new cannot be a part of FB
                    remove Sol_i from W (excretion)
                    insert a random Sol into W to replace Sol_i
                endif
                Sol_new is reabsorbed
            else
                if it is better than the Sol_worst in FB
                    Sol_worst is secreted
                else
                    Sol_i is secreted

                endif
            endif
        endfor
        evaluate the GFR of solutes in FB
        if GFR<15
            if this is the second kidney with GFR<15
                if dialysis is on
                    do the KA process with population in FB
                else
                    transplant is applied
                endif
            else
                send all solutes from current kidney to another kidney
            endif
        endif
        if 15<GFR<60
            apply movement of solutes in FB
        endif
    rank the Sols from FB and update the Sol_best
    merge W and FB
    update fr, Eq. 2
    endfor
update overall Sol_best
end while
return overall Sol_best
```

**FIGURE 5.** Pseudocode of dual kidney-inspired algorithm.

dialysis machine and, again, more exploration is the outcome of that process. A transplant involves restarting the population with random solutions to again provide the algorithm with more exploration. Overall, the goal of all these proposed modifications is to enhance the performance of the algorithm and to help it find a better solution during its search of the search space. Brief explanations of the operations in Dual-KA are:

*Filtration*: separation of solutions and them is being sent into FB and W using Eq. 2 as a threshold.

*Reabsorption*: shifting a potential new member of W to FB, if it meets the condition for inclusion in FB.

*Secretion*: secreting from FB when a new element of FB is not better than the worst solution in FB.

*Excretion*: removing a solution from W if no success in reabsorption operation.

*Treatment*: applying Eq. 1 to the solutions in FB

*Dialysis*: performing the KA process on the FB population.

*Transplant*: restarting one of the sub-populations with random solutions.

The pseudocode of Dual-KA is given in Fig. 5 and the overall cooperation between two KAs is illustrated in Fig. 6.

## IV. EXPERIMENTAL RESULTS

Two experiments are conducted to evaluate the performance of Dual-KA and the results are compared with KA and other available algorithms in the literature. First, 12 test functions are used to test the performance of the algorithm.

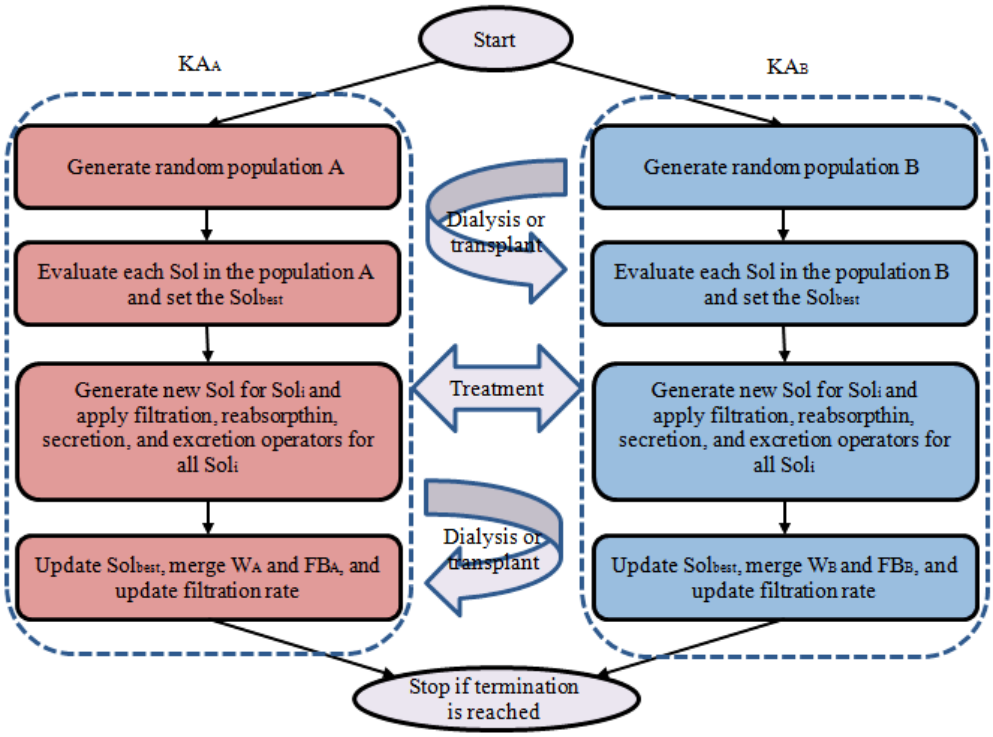


**FIGURE 6.** Overall cooperation between two KAs in Dual-KA.

In the second experiment, benchmark classification and time series prediction problems are used to further evaluate the algorithm.

### A. RESULTS FOR TEST FUNCTIONS

The 12 test functions on which KA and Dual-KA are tested are detailed below in Table 1. The dimensional size for all test functions is set to 100 (d=100). In this experiment, the size of the population is set to 100, as well as the number of iterations. The mean of the results of 100 runs and the standard deviation for each test function are presented in Table 2. The optimal solution is also given in Table 2 so that it can be compared with the results obtained by the algorithm.

Table 2 shows that Dual-KA can find the global optimum in all but one of the test functions, while KA was not successful in three out of 12 test functions. This proves the significant difference between KA and Dual-KA. The ability of Dual-KA to find the global optimum for 11 out of 12 test functions is due to the greater amount of exploration provided by cooperation between kidneys and more efforts made to find the global optimum through the use of the dialysis and transplant process embedded in this algorithm. This enables the algorithm to search more areas in the search space and find a better solution that could perhaps be the global optimum. A detailed comparison of the convergence capability and optimization progress of KA and Dual-KA is provided in Fig. 7. From Fig. 7 it can be observed that most of the time during the optimization process Dual-KA can discover the optimal solution in earlier steps compared to KA. This is proof of the potential optimization ability of Dual-KA when it is designed to make use of the cooperation that exists between two kidneys.

### B. RESULTS FOR BENCHMARK CLASSIFICATION AND TIME SERIES PREDICTION PROBLEMS

Some standard classification problems, along with time series prediction problems are employed to test and compare the optimization ability of KA and Dual-KA. Table 3 provides the details of the datasets that are used in this experiment; the first six datasets are classification datasets and the last two are time series prediction datasets. The algorithms were run 20 times. The classification datasets are taken from the UCI machine learning repository [35]. The data for the Mackey-Glass time series dataset is calculated from an equation specified in [4]. The Gas Furnace time series dataset can be collected at http://datasets.connectmv.com/datasets. Because feature selection is important for classification problem and affects classification accuracy, in the case of classification problems in the proposed method, selection of features is randomly performed and ANN is structured based on selected features in each solution construction. Therefore once an ANN model is optimized, selection of features is also optimized. To do this a dynamic vector represents the solution where the number of cells in this vector is changed due to changing the number of selected features in each solution construction.

**TABLE 1.** Details of the test functions.

| Test function | Equation for test function | Global optimum | Dimension |
|---|---|---|---|
| Michalewicz function | $f(x) = -\sum_{i=1}^{d} \sin(x_i) \sin^{2m}\left(\frac{ix_i^2}{\pi}\right)$ | $f(x)$ = -1.8013<br>$x$ =(2.20, 1.57) | $0 < x < \pi$ |
| Rosenbrock function | $f(x) = \sum_{i=1}^{d-1} \left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right]$ | $f(x)$ = 0<br>$x$ = (1, 1) | -2.048 < $x$ < 2.048 |
| De Jong function | $f(x) = \sum_{i=1}^{d} x_i^2$ | $f(x)$ = 0<br>$x$ = (0, 0) | -5.12 < $x$ < 5.12 |
| Schwefel function | $f(x) = 418.9829d - \sum_{i=1}^{d} x_i \sin(\sqrt{\lvert x_i \rvert})$ | $f(x)$ = 0<br>$x$ = (420.9687, 420.9687) | -500 < $x$ < 500 |
| Ackley function | $f(x) = -20\exp(-0.2\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2}) - \exp(\frac{1}{d}\sum_{i=1}^{d}\cos(2\pi x_i)) + 20 + e$ | $f(x)$ = 0<br>$x$ = (0, 0) | -32.768 < $x$ <32.768 |
| Rastrigin function | $f(x) = 10d + \sum_{i=1}^{d} \left[x_i^2 - 10\cos(2\pi x_i)\right]$ | $f(x)$ = 0<br>$x$ = (0, 0) | -5.12 < $x$ < 5.12 |
| Easom function | $f(x) = -\cos(x_1)\cos(x_2)\exp(-(x_1 - \pi)^2 - (x_2 - \pi)^2)$ | $f(x)$ = -1<br>$x$ = ($\pi$, $\pi$) | -100 < $x$ <100 |
| Griewangk function | $f(x) = \sum_{i=1}^{d} \frac{x_i^2}{4000} - \prod_{i=1}^{d} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$ | $f(x)$ = 0<br>$x$ = (0, 0) | -600 < $x$ < 600 |
| Shifted sphere function | $f(x) = \sum_{i=1}^{D} z_{ij}^2 + f\_bias$, z = $x$ − o , x= [$x_1$ ,$x_2$ ,... ,$x_D$ ] | $f(x)$ = -450<br>$x$=$o$ | -100 < x < 100 |
| Shifted rosenbrock's function | $f(x) = \sum_{i=1}^{D-1}(100(z_i^2 - z_{i+1})^2 + (z_i - 1)^2) + f\_bias$, $z$=$x$-$o$+1, x= [$x_1$ ,$x_2$ ,... ,$x_D$ ] | $f(x)$ = 390<br>$x$=$o$ | -100 < x < 100 |
| Shifted rotated rastrigin's function | $f(x) = \sum_{i=1}^{D}(z_i^2 - 10\cos(2\pi z_i) + 10 + f\_bias$ , $z$=($x$-$o$)*$M$, x= [$x_1$ ,$x_2$ ,... ,$x_D$ ] | $f(x)$ = -330<br>$x$=$o$ | -5 < x < 5 |
| Rotated hyper-ellipsoid function | $f(x) = \sum_{i=1}^{d}\sum_{j=1}^{i} x_j^2$ | $f(x)$ = 0<br>$x$ = (0, 0) | -65.536 < $x$ < 5.536 |

**TABLE 2.** Results for the test functions.

| | | KA | | Dual-KA | |
|---|---|---|---|---|---|
| **Test function** | **Optimum** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** |
| Michalewicz function | -1.8013 | -1.8142 | 3.20E-04 | -1.8183 | 1.30E-03 |
| Rosenbrock function | 0 | 0 | 0 | 0 | 0 |
| De Jong function | 0 | 0 | 0 | 0 | 0 |
| Schwefel function | 0 | 1.83E-06 | 1.430E-07 | 0 | 0 |
| Ackley function | 0 | 0 | 0 | 0 | 0 |
| Rastrigin function | 0 | 0 | 0 | 0 | 0 |
| Easom function | -1 | -1 | 0 | -1 | 0 |
| Griewangk function | 0 | 0 | 0 | 0 | 0 |
| Shifted sphere function | -450 | -450 | 0 | -450 | 0 |
| Shifted rosenbrock's function | 390 | 388.3204 | 5,36E-03 | 390 | 0 |
| Shifted rotated rastrigin's function | -330 | -330 | 0 | -330 | 0 |
| Rotated hyper-ellipsoid function | 0 | 0 | 0 | 0 | 0 |

The ANN used for this experiment consists of two hidden layers and two nodes for each hidden layer. This structure has been used successfully in the literature [36]–[38]. The hyperbolic tangent is the best activation function for an ANN [39] and is the activation function used for this test. Each solution is represented by a one-dimensional vector, where the weights and biases of the ANN are located in each cell of the vector.

For the classification problems, the conditional attributes are considered as the input of the ANN and the class attribute as the output. The number of inputs is equal to the number of

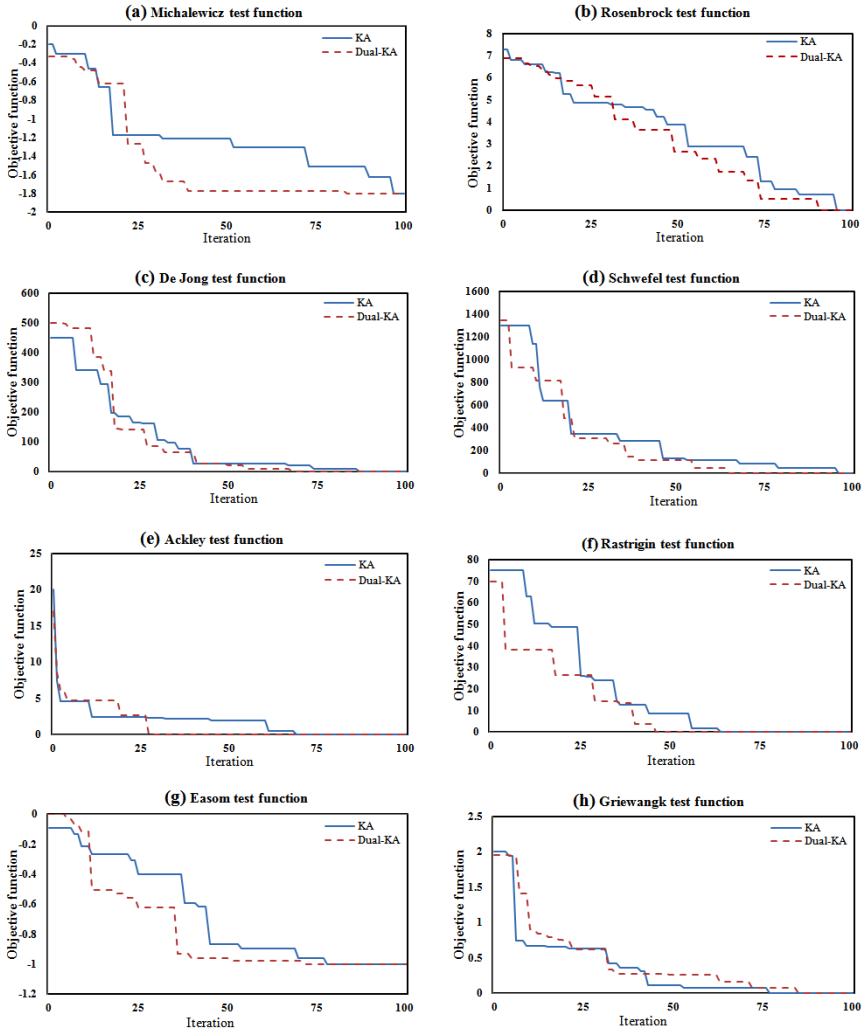


**FIGURE 7.** Optimization process of KA and Dual-Ka for: (a) Michalewicz, (b) Rosenbrock, (c) De Jong, (d) Schwefel, (e) Ackley, (f) Rastrigin, (g) Easom and (h) Griewangk test functions.

**TABLE 3.** Characteristics of datasets.

| Dataset | No of examples | No of features | No of classes |
|---|---|---|---|
| Iris | 150 | 4 | 3 |
| Diabetes | 768 | 8 | 2 |
| Thyroid | 7,200 | 21 | 3 |
| Cancer | 699 | 10 | 2 |
| Card | 690 | 15 | 2 |
| Glass | 214 | 10 | 6 |
| Mackey-Glass | 1,000 | 1 | 0 |
| Gas Furnace | 296 | 2 | 0 |

conditional attributes. The hidden layers consist of two nodes in each layer and the output layer has one node for the class attribute. For the Mackey-Glass time series prediction dataset $x(t), x(t-6), x(t-12)$, and $x(t-18)$ are the input (four input node) and $x(t+6)$ is the output for the ANN. The output for the Gas Furnace time series prediction dataset is $y(t)$ and the inputs with six nodes are $u(t-3)$, $u(t-2)$, $u(t-1)$, $y(t-3)$, $y(t-2)$, and $y(t-1)$. The selection of inputs and outputs is the same as in earlier works [36], [38]. Similar to previous research [4], [37], the data was normalized into the range of [0, 1] for the ANN in order to measure the performance of the algorithms. We used 10-fold cross validation to estimate the performance of the model. The data were randomly divided into two parts for each run. One half was employed for the training set and the other half was used as the testing set to test the final model. The overall process conducted

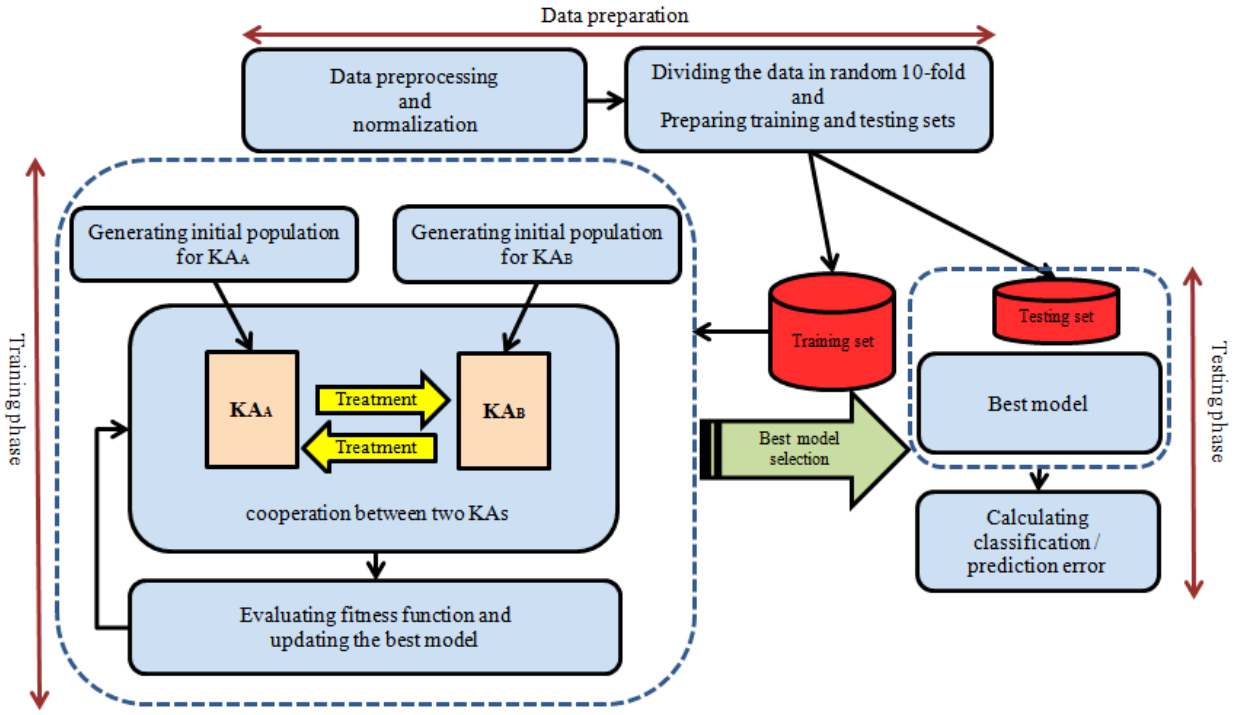


**FIGURE 8.** Overall process conducted for classification and time series prediction.

for classification and time series prediction is illustrated in Fig. 8.

The results of applying KA and Dual-KA to standard classification problems and afterwards time series prediction problems are given in Table 4.

For classification problems the training classification error and testing classification error are calculated for the training set and testing set respectively. For the Mackey-Glass dataset, the root mean square error is calculated for the training set and testing set and for the Gas Furnace dataset the mean square error is used. From Table 4 it is clearly observable that in all cases Dual-KA provides better results compared to KA. This superiority is the result of the cooperation between the two kidneys and the dialysis and transplant process in this algorithm which enables the algorithm to undertake more exploration of the search space and leads the algorithm towards a better solution.

A t-test analysis for training and testing errors is also performed to compare the performance of KA and Dual-KA. The $p$-values for all the datasets are calculated where the critical value $\alpha$ for this experiment is equal to 0.05. Values lower than the critical value show that there is an improvement in the result provided by Dual-KA. These improvements are shown in bold in Table 5. The extremely small $p$-values in Table 5 indicate that Dual-KA has better ability compared to KA. These results support those presented in Table 4.

The optimization ability of Dual-KA, when applied to classification and time series prediction problems, is further compared with that of the following methods in the literature: (1) Kidney-inspired algorithm with treatment process (KAT) [33], (2) kidney-inspired algorithm with control of filtration rate (Log-*aCtrl-KA)* [28], (3) a multi-population cuckoo search algorithm with control of parameter $P_a$ (Multipop-$P_a$Ctrl-CS) [22], (4) a cooperative-competitive global best harmony search algorithm (CC-GHS) [3], (5) a multi-population bat algorithm (BatRM-S) [4], (6) a Taguchi-based logistic bat algorithm (T-LogisticBatDNN) [38], and (7) an integration of tabu search, simulated annealing, and GA with backpropagation (for classification only) (GaTSa+BP) [37]. Twenty runs are carried out and the results of averages of 20 runs for the classification and time series prediction problems are provided in Table 6 and Table 7, respectively. The best results are presented in bold in both tables for ease of comparison.

Dual-KA achieved the best result in diabetes data for classification. Although Dual-KA cannot attain the best results, in all datasets, compared to the other methods in the literature, in overall, it shows comparable results with the best method in both classification and time series prediction problems. If we look at the KAT and Log-*aCtrl-KA* results and compare it with Dual-KA and other methods, it can be seen that even these two recently published methods cannot achieve the best results except in cancer data for Log-*aCtrl-KA* method. The results of Dual-KA are a promising finding, motivating the researchers to undertake further improvements in future work.

A statistical analysis of the results is also conducted using a Friedman test in order to ascertain whether there are any

**TABLE 4.** Classification and time series prediction results obtained by KA and Dual-KA.

| Dataset | Criteria | KA | Dual-KA |
|---|---|---|---|
| Iris | Training error % | 2.4213 | **1.5877** |
| | Std. Dev. | 0.0547 | 0.0543 |
| | Testing error% | 2.6738 | **1.6242** |
| | Std. Dev. | 0.0726 | 0.0567 |
| Diabetes | Training error % | 19.1576 | **12.9970** |
| | Std. Dev. | 0.0028 | 0.0032 |
| | Testing error% | 20.8243 | **14.9918** |
| | Std. Dev. | 0.0625 | 0.0734 |
| Thyroid | Training error % | 6.7401 | **4.8463** |
| | Std. Dev. | 0.0172 | 0.0246 |
| | Testing error% | 7.6212 | **5.1183** |
| | Std. Dev. | 0.0726 | 0.0674 |
| Cancer | Training error % | 2.9947 | **2.7825** |
| | Std. Dev. | 0.0762 | 0.0565 |
| | Testing error% | 3.2514 | **2.8329** |
| | Std. Dev. | 0.0289 | 0.0385 |
| Card | Training error % | 14.6352 | **12.6200** |
| | Std. Dev. | 0.0459 | 0.0293 |
| | Testing error% | 16.4859 | **13.1415** |
| | Std. Dev. | 0.0027 | 0.00395 |
| Glass | Training error % | 39.7632 | **29.4850** |
| | Std. Dev. | 0.0043 | 0.0036 |
| | Testing error% | 40.4299 | **31.4364** |
| | Std. Dev. | 0.0836 | 0.0621 |
| Mackey-Glass | Training error % | 3.20E-03 | **1.54E-03** |
| | Std. Dev. | 0.0001 | 0.0003 |
| | Testing error% | 3.47E-03 | **1.71E-03** |
| | Std. Dev. | 4.60E-06 | 3.40E-06 |
| Gas Furnace | Training error % | 0.3941 | **0.2317** |
| | Std. Dev. | 5.00E-06 | 3.00E-08 |
| | Testing error% | 0.4426 | **0.2598** |
| | Std. Dev. | 0.0072 | 0.0033 |

**TABLE 5.** Pairwise comparison of *p*-values.

| Dataset | KA | |
|---|---|---|
| | Training error | Testing error |
| Iris | **6.06E-06** | **0.0003** |
| Diabetes | **9.40E-05** | **2.35E-05** |
| Thyroid | **0.0084** | **5.68E-06** |
| Cancer | 0.1564 | **0.0033** |
| Card | **0.0099** | **0.0011** |
| Glass | **0.0003** | **8.02E-05** |
| Mackey-Glass | **8.92E-06** | **2.25E-06** |
| Gas Furnace | **0.0001** | **2.01E-05** |

**TABLE 6.** Comparison of Dual-KA with other methods in literature for classification problems.

| Dataset | Dual-KA | KAT[33] | Log-*aCtrl-KA* [28] | Multipop-$P_a$Ctrl-CS [22] | CC-GHS [3] | BatRM-S [4] | T-LogisticBatDNN [38] | GaTSa+BP [37] |
|---|---|---|---|---|---|---|---|---|
| Iris | 1.6242 | 1.6575 | 1.7621 | 1.6906 | 1.9573 | **1.5309** | 1.8313 | 5.2564 |
| Diabetes | **14.9918** | 15.3251 | 21.2831 | 15.6356 | 16.6356 | 19.3021 | 20.139 | 27.061 |
| Thyroid | 5.1183 | 5.1436 | 6.0273 | **5.1104** | 5.4104 | 6.2435 | 6.7545 | 7.1509 |
| Cancer | 2.8329 | 2.8996 | **2.2733** | 2.4596 | 3.0263 | 2.9928 | 2.8947 | 7.192 |
| Card | 13.1415 | 13.0705 | 14.3479 | **12.5831** | 13.1498 | 13.4163 | 13.738 | 15.242 |
| Glass | 31.4364 | 31.6766 | 37.5217 | **28.0919** | 34.0919 | 38.4251 | 38.429 | 55.142 |

big differences between the methods compared in this experiment. For the classification problems the Friedman test result is 29.0555, which is greater than the critical value (10.57), so the null hypothesis can be rejected and it can be concluded that there is a major difference in the performance of the tested algorithms. A Nemenyi post-hoc test is also conducted

**TABLE 7.** Comparison of Dual-KA with other methods in literature for Time series prediction problems.

| Dataset | Dual-KA | KAT[33] | Log-*aCtrl-KA*[28] | Multipop-$P_a$Ctrl-CS [22] | CC-GHS [3] | BatRM-S [4] | T-LogisticBatDNN [38] |
|---|---|---|---|---|---|---|---|
| Mackey-Glass | 1.71E-03 | 1.70E-03 | 0.016 | 0.0014 | **1.00E-05** | 0.0019 | 0.0027 |
| Gas Furnace | 0.2598 | 0.2664 | 0.4731 | **0.1903** | 0.287 | 0.5518 | 0.33 |

**TABLE 8.** Nemenyi test results for the classification problems.

| Algorithms | | Dual-KA | KAT | Log-*aCtrl-KA* | Multipop-$P_a$Ctrl-CS | CC-GHS | BatRM-S | T-LogisticBatDNN |
|---|---|---|---|---|---|---|---|---|
| | Mean | 11.52423 | 11.62882 | 13.86923 | 10.92853 | 12.37855 | 13.65178 | 13.96442 |
| Dual-KA | 11.52423 | - | - | **2.345008** | 0.595692 | **0.854324** | **2.127558** | **2.440191** |
| KAT | 11.62882 | - | - | **2.240417** | **0.700283** | **0.749733** | **2.022967** | **2.3356** |
| Log-*aCtrl-KA* | 13.86923 | - | - | - | **2.9407** | **1.490683** | **0.21745** | 0.095183 |
| Multipop-$P_a$Ctrl-CS | 10.92853 | - | - | - | - | **1.450017** | **2.72325** | **3.035883** |
| CC-GHS | 12.37855 | - | - | - | | | **1.273233** | **1.585867** |
| BatRM-S | 13.65178 | - | - | - | - | - | - | **0.312633** |
| T-LogisticBatDNN | 13.96442 | - | - | - | - | - | - | - |

to find out where the differences lie. The results of this test are presented in Table 8. The Minimum Significant Difference (MSD) in this test is equal to 0.6480. Values greater than the MSD indicate that there are statistically major differences; these values are presented in bold in Table 8.

The Friedman test is also performed for the time series prediction problems. In this case, the $p$-value for the Friedman test is equal to 6.4 which is less than the critical value of 7.6. Therefore, the null hypothesis cannot be rejected. There is not a big difference between the results obtained by the compared methods for time series prediction. Therefore, any kind of post-hoc test is not applicable in this analysis. Hence, although there are differences among the results presented in Table 7, according to the results of the statistical analysis, these differences are not significant.

### C. RESULTS FOR REAL-WORLD PROBLEMS

In this section two problems discussed in this paper (classification and time series prediction), were considered from real life in order to test the ability of the proposed method to solve real-world problems. In the first following subsection, prediction of water quality from real-world data is discussed and in the second subsection, two medical datasets, leukemia and prostate, were classified for cancer detection.

#### 1) REAL-WORLD WATER QUALITY PREDICTION PROBLEM

The performance of Dual-KA is also tested on real-world water quality prediction data. The data comes from a weather station near Kajang in the state of Selangor in Malaysia. The algorithms are applied to monthly water data for the years 2004 to 2013. This data is multivariate and it has 13 attributes: SFLOW, TEMP (degrees C), TUR (NTU), DS (mg/l), TS (mg/l), $NO_3$(mg/l), $PO_4$(mg/l), DO (mg/l), BOD (mg/l), COD (mg/l), SS (mg/l), pH (unit), and $NH_3$-NL (mg/l). All the attributes are used for the ANN and the last six attributes, which are the most important attributes for water quality prediction, are employed as the outputs of the ANN. Seventy percent of the data is used for the training set and thirty percent for the testing set. The data is normalized into a range of between 0 and 1 by using the min-max normalization technique so it can be matched with the ANN. A 10-fold cross-validation is used to validate the results. Twenty runs are carried out and a one-step-ahead prediction is performed. The average of 20 runs is presented in Table 9.

**TABLE 9.** Prediction results of Dual-KA for real-world water quality data.

| Feature | Criteria | Dual-KA |
|---|---|---|
| $NH_3$-NL | Training error% | 0.1995 |
| | Testing error% | 0.2168 |
| PH | Training error% | 0.0946 |
| | Testing error% | 0.0990 |
| SS | Training error% | 0.1245 |
| | Testing error% | 0.1258 |
| COD | Training error% | 0.2481 |
| | Testing error% | 0.2759 |
| BOD | Training error% | 0.1446 |
| | Testing error% | 0.1448 |
| DO | Training error% | 0.1053 |
| | Testing error% | 0.1469 |

Fig. 9 shows a sample graph of the actual value (solid line) and model output (dotted line) of Dual-KA for the attribute NH3-NL.

The performance of Dual-KA when applied to the real-world problem of water quality prediction is also compared with a method proposed in the literature,

**TABLE 10.** Comparison of Dual-KA and Multipop-$P_a$ Ctrl-CS for water quality prediction problem.

| Feature | Dual-KA | Multipop-$P_a$Ctrl-CS [22] |
|---|---|---|
| $NH_3$-NL | 0.2168 | 0.2169 |
| PH | 0.0990 | 0.1169 |
| SS | 0.1258 | 0.1609 |
| COD | 0.2759 | 0.3169 |
| BOD | 0.1448 | 0.1166 |
| DO | 0.1469 | 0.1569 |

**TABLE 11.** Comparison of Dual-KA results for classification accuracy and other methods in cancer detection.

| Method | Leukemia | Prostate |
|---|---|---|
| Hernandez [41] | 91.5% | N/A |
| Glaab [42] | N/A | 94% |
| Yu [43] | 95.55% | N/A |
| Hengpraprohm [44] | 91.9% | N/A |
| Gunavathi [45] | N/A | 85.71% |
| Nguyen [46] | 96.48% | 95.60% |
| Salem [47] | 97.06% | N/A |
| Esmaeili [40] | 98.57% | 96.77% |
| Dual-KA | 99.34% | 96.98% |

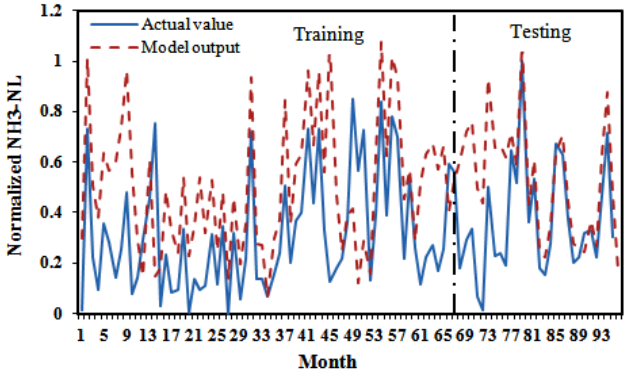


**FIGURE 9.** Comparison of the actual value and the model output of Dual-KA for $NH_3$-Nl.

Multipop-$P_a$Ctrl-CS [22]. The results of testing error are presented in Table 10, in which the results show that Dual-KA has higher ability in most of the cases.

### 2) REAL-WORLD CANCER DETECTION PROBLEM

The proposed method was also tested on two other datasets, leukemia and prostate (genomic data), in order to find ability and performance of the proposed method when applied to other real-worlds problems. The data in leukemia dataset is categorised into two classes of ALL and AML with 7129 genes and 72 samples. The prostate dataset with 12,600 genes and 102 samples contains two classes of normal and cancerous. The results are compared and evaluated with other methods available in the literature in Table 11. The gene selection was performed using selected genes reported in [40]. The classification accuracy of leukemia data by the proposed method was increased to 99.34%. In addition, the classification accuracy, for prostate data came to 96.98%. The results of both leukemia and prostate data show enhancement compared to others in this experiment (see Table 11). This is proof of the ability of the Dual-KA for real-world classification problems.

## V. CONCLUSION

In this paper, Dual-KA was developed, which is a multi-population approach stimulated by the cooperation between the two kidneys in the human body. In Dual-KA, the GFR is checked in each iteration of each population. If the GFR is greater than 60, the process is continued as normal. If it is between 15 and 60, a kind of treatment is performed. However, if the GFR is less than 15 and the current sub-population is the first sub-population with a GFR of less than 15 in the current iteration, all the solutions in the current sub-population are sent to the other sub-population. If the second kidney fails, a dialysis or transplant process is carried out. After sending the solutions to the other sub-population, the sub-population with all the solutions performs the process with a new filtration rate. The aim of this contribution is to enable the algorithm to better explore the search space. The proposed method was tested on 12 test functions, some benchmark classification and time series prediction problems, and real-world water quality prediction and cancer detection problems. The results showed that Dual-KA performed better than basic KA, but, in comparison with the best-performing methods in the literature, it showed comparable results. The proposed method is applicable for any problem in real-world and, as future work; ability of the proposed method will be tested on genomic data in medicine.

## REFERENCES

[1] H.-Z. Li, S. Guo, C.-J. Li, and J.-Q. Sun, "A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm," *Knowl.-Based Syst.*, vol. 37, pp. 378–387, Jan. 2013.

[2] S. Ding, C. Su, and J. Yu, "An optimizing BP neural network algorithm based on genetic algorithm," *Artif. Intell. Rev.*, vol. 36, no. 2, pp. 153–162, Aug. 2011.

[3] N. S. Jaddi and S. Abdullah, "A cooperative-competitive master-slave global-best harmony search for ANN optimization and water-quality prediction," *Appl. Soft Comput.*, vol. 51, pp. 209–224, Feb. 2017.
[4] N. S. Jaddi, S. Abdullah, and A. R. Hamdan, "Multi-population cooperative bat algorithm-based optimization of artificial neural network model," *Inf. Sci.*, vol. 294, pp. 628–644, Feb. 2015.
[5] W. Zhang, G. Han, J. Wang, and Y. Liu, "A BP neural network prediction model based on dynamic cuckoo search optimization algorithm for industrial equipment fault prediction," *IEEE Access*, vol. 7, pp. 11736–11746, 2019.
[6] A. A. Heidari, H. Faris, I. Aljarah, and S. Mirjalili, "An efficient hybrid multilayer perceptron neural network with grasshopper optimization," *Soft Comput.*, vol. 23, no. 17, pp. 7941–7958, Sep. 2019.
[7] M. W. Guo, J. S. Wang, L. F. Zhu, S. S. Guo, and W. Xie, "An improved grey wolf optimizer based on tracking and seeking modes to solve function optimization problems," *IEEE Access*, vol. 8, pp. 69861–69893, 2020.
[8] N. S. Jaddi and M. Saniee Abadeh, "DNA methylation-based age prediction using cell separation algorithm," *Comput. Biol. Med.*, vol. 121, Jun. 2020, Art. no. 103747.
[9] N. S. Jaddi and S. Abdullah, "Global search in single-solution-based metaheuristics," *Data Technol. Appl.*, 2020, doi: 10.1108/DTA-07-2019-0115.
[10] F. vandenBergh and A. P. Engelbrecht, "A cooperative approach to particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 225–239, Jun. 2004.
[11] B. Niu, Y. Zhu, and X. He, "Multi-population cooperative particle swarm optimization," in *Advances in Artificial Life*. Berlin, Germany: Springer, 2005, pp. 874–883.
[12] B. Yu, B. Jiao, and X. Gu, "Cooperative particle swarm optimizer based on multi-population and its application to flow-shop scheduling problem," in *Proc. Asia Simulation Conf. 7th Int. Conf. Syst. Simulation Sci. Comput.*, Oct. 2008, pp. 1536–1542.
[13] C.-K. Goh and K. Chen Tan, "A competitive-cooperative coevolutionary paradigm for dynamic multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 1, pp. 103–127, Feb. 2009.
[14] B. Niu, Y. Zhu, X. He, and H. Wu, "MCPSO: A multi-swarm cooperative particle swarm optimizer," *Appl. Math. Comput.*, vol. 185, no. 2, pp. 1050–1062, Feb. 2007.
[15] J. Togelius, P. Burrow, and S. M. Lucas, "Multi-population competitive co-evolution of car racing controllers," in *Proc. IEEE Congr. Evol. Comput.*, Sep. 2007, pp. 4043–4050.
[16] C. K. Goh, K. C. Tan, D. S. Liu, and S. C. Chiam, "A competitive and cooperative co-evolutionary approach to multi-objective particle swarm optimization algorithm design," *Eur. J. Oper. Res.*, vol. 202, no. 1, pp. 42–54, 2010.
[17] L. Zhao and Y. Yang, "PSO-based single multiplicative neuron model for time series prediction," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2805–2812, Mar. 2009.
[18] B. Li, T.-S. Lin, L. Liao, and C. Fan, "Genetic algorithm based on multi-population competitive coevolution," in *Proc. IEEE Congr. Evol. Comput., IEEE World Congr. Comput. Intell.*, Jun. 2008, pp. 225–228.
[19] Y.-N. Guo and D. Liu, "Multi-population cooperative particle swarm cultural algorithms," in *Proc. 7th Int. Conf. Natural Comput.*, Jul. 2011, pp. 199–206.
[20] K. C. Tan, Y. H. Chew, T. H. Lee, and Y. J. Yang, "A cooperative coevolutionary algorithm for multiobjective optimization," in *Proc. Conf. IEEE Int. Conf. Syst., Man Cybern. Conf. Theme Syst. Secur. Assurance (SMC)*, Oct. 2004, pp. 1926–1931.
[21] K. C. Tan, Y. J. Yang, and C. K. Goh, "A distributed cooperative coevolutionary algorithm for multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 10, no. 5, pp. 527–549, Oct. 2006.
[22] N. S. Jaddi, S. Abdullah, and M. A. Malek, "Master-Leader-Slave cuckoo search with parameter control for ANN optimization and its real-world application to water quality prediction," *PLoS ONE*, vol. 12, no. 1, Jan. 2017, Art. no. e0170372.
[23] N. S. Jaddi, J. Alvankarian, and S. Abdullah, "Kidney-inspired algorithm for optimization problems," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 42, pp. 358–369, Jan. 2017.
[24] M. Ehteram, H. Karami, S. F. Mousavi, S. Farzin, A. B. Celeste, and A.-E. Shafie, "Reservoir operation by a new evolutionary algorithm: Kidney algorithm," *Water Resour. Manage.*, vol. 32, pp. 4681-4706, Nov. 2018.
[25] S. Ekinci, A. Demiroren, and B. Hekimoglu, "Parameter optimization of power system stabilizers via kidney-inspired algorithm," *Trans. Inst. Meas. Control*, vol. 41, no. 5, pp. 1405–1417, 2019.
[26] N. N. Farahani, S. Farzin, and H. Karami, "Flood routing by Kidney algorithm and Muskingum model," *Natural Hazards*, pp. 1–19, Sep. 2018, doi: 10.1007/s11069-018-3482-x.
[27] A. A. B. Homaid, A. A. Alsewari, A. K. Alazzawi, and K. Z. Zamli, "A kidney algorithm for pairwise test suite generation," *Adv. Sci. Lett.*, vol. 24, no. 10, pp. 7284–7289, Oct. 2018.
[28] N. S. Jaddi and S. Abdullah, "Optimization of neural network using kidney-inspired algorithm with control of filtration rate and chaotic map for real-world rainfall forecasting," *Eng. Appl. Artif. Intell.*, vol. 67, pp. 246–259, Jan. 2018.
[29] Q. Ke and J. Wen, "Kidney-inspired algorithm based on scaling factor and cooperative operator," *Comput. Sci. Appl.*, vol. 8, no. 4, pp. 472–479, 2018.
[30] M. K. Taqi and R. Ali, "OBKA-FS: An oppositional-based binary kidneyinspired search algorithm for feature selection," *J. Theor. Appl. Inf. Technol.*, vol. 95, no. 1, p. 9, 2017.
[31] S. Ekinci, A. Demirören, H. L. Zeynelgil, and S. Kaya, "Design of PID controller for automatic voltage regulator system through kidney-inspired algorithm," *Fen Bilimleri Dergisi*, vol. 7, no. 2, pp. 383–398, 2019.
[32] S. Ekinci and B. Hekimoğlu, "Improved kidney-inspired algorithm approach for tuning of PID controller in AVR system," *IEEE Access*, vol. 7, pp. 39935–39947, 2019.
[33] N. S. Jaddi and S. Abdullah, "Kidney-inspired algorithm with reduced functionality treatment for classification and time series prediction," *PLoS ONE*, vol. 14, no. 1, Jan. 2019, Art. no. e0208308.
[34] J. E. Hall, *Guyton and Hall Textbook of Medical Physiology E-Book*. Amsterdam, The Netherlands: Elsevier, 2015.
[35] C. L. Blake and C. J. Merz. (1998). *UCI Repository of Machine Learning Databases*. [Online]. Available: http://www.ics.uci.edu/~mlearn/
[36] S.-K. Oh, W. Pedrycz, and S.-B. Roh, "Hybrid fuzzy set-based polynomial neural networks and their development with the aid of genetic optimization and information granulation," *Appl. Soft Comput.*, vol. 9, no. 3, pp. 1068–1089, Jun. 2009.
[37] C. Zanchettin, T. B. Ludermir, and L. M. Almeida, "Hybrid training method for MLP: Optimization of architecture and training," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 4, pp. 1097–1109, Aug. 2011.
[38] N. S. Jaddi, S. Abdullah, and A. R. Hamdan, "Optimization of neural network model using modified bat-inspired algorithm," *Appl. Soft Comput.*, vol. 37, pp. 71–86, Dec. 2015.
[39] B. Karlik and A. Vehbi, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," *Int. J. Artif. Intell. Expert Syst.*, vol. 1, no. 4, pp. 111–122, 2015.
[40] R. E. Naftchali and M. S. Abadeh, "A multi-layered incremental feature selection algorithm for adjuvant chemotherapy effectiveness/futileness assessment in non-small cell lung cancer," *Biocybernetics Biomed. Eng.*, vol. 37, no. 3, pp. 477–488, 2017.
[41] J. C. H. Hernandez, B. Duval, and J.-K. Hao, "A genetic embedded approach for gene selection and classification of microarray data," in *Proc. Eur. Conf. Evol. Comput., Mach. Learn. Data Mining Bioinf.*, vol. 2007, pp. 90–101.
[42] E. Glaab, J. Bacardit, J. M. Garibaldi, and N. Krasnogor, "Using rule-based machine learning for candidate disease gene prioritization and sample classification of cancer gene expression data," *PLoS ONE*, vol. 7, no. 7, Jul. 2012, Art. no. e39932.
[43] H. Yu, J. Ni, Y. Dan, and S. Xu, "Mining and integrating reliable decision rules for imbalanced cancer gene expression data sets," *Tsinghua Sci. Technol.*, vol. 17, no. 6, pp. 666–673, Dec. 2012.
[44] S. Hengpraprohm, "GA-based classifier with SNR weighted features for cancer microarray data classification," *Int. J. Signal Process. Syst.*, vol. 1, no. 1, pp. 29–33, 2013.
[45] C. Gunavathi and K. Premalatha, "Performance analysis of genetic algorithm with kNN and SVM for feature selection in tumor classification," *Int. J. Comput. Electron Autom. Control Inf. Eng.*, vol. 8, no. 8, pp. 1490–1497, 2014.
[46] T. Nguyen, A. Khosravi, D. Creighton, and S. Nahavandi, "Hidden Markov models for cancer classification using gene expression profiles," *Inf. Sci.*, vol. 316, pp. 293–307, Sep. 2015.
[47] H. Salem, G. Attiya, and N. El-Fishawy, "Classification of human cancer diseases by gene expression profiles," *Appl. Soft Comput.*, vol. 50, pp. 124–134, Jan. 2017.

**SALWANI ABDULLAH** received the B.Sc. degree in computer science from the Universiti Teknologi Malaysia, the master's degree specializing in computer science from the Universiti Kebangsaan Malaysia (UKM), and the Ph.D. degree in computer science from the University of Nottingham, U.K. She is currently a Professor of computational optimization with the Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia. Her research interests include artificial intelligence and operation research, particularly computational optimization algorithms (heuristic and meta-heuristic, evolutionary algorithms, local search) that involve different real world applications for single and multi-objective continuous and combinatorial optimization problems such as timetabling, scheduling, routing, nurse rostering, dynamic optimization, data mining problems (feature selection, clustering, classification, time-series prediction), intrusion detection, and search-based software testing.

**NAJMEH SADAT JADDI** received the B.S. degree in software engineering from Staffordshire University, (Malaysia campus), Stoke-on-Trent, U.K., in 2008, and the M.S. and Ph.D. degrees in information technology from The National University of Malaysia, in 2011 and 2014, respectively. She joined the Faculty of Information Science and Technology, The National University of Malaysia, in 2008. Her current research interests include data mining, artificial intelligence, machine learning, artificial neural networks, deep learning, optimization algorithms and applications, and mathematical modeling.