IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 2, 2018, accepted December 14, 2018, date of publication December 21, 2018,
date of current version February 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2889151

A New Evolutionary Machine Learning Approach for Identifying Pyrene Induced Hepatotoxicity and Renal Dysfunction in Rats

JIAYIN ZHU[1], FENGTING ZHU[2], SIYI HUANG[3], GANG WANG[4], HUILING CHEN[5], XUEHUA ZHAO[6], AND SHU-YUN ZHANG[7]

[1]Laboratory Animal Center, Wenzhou Medical University, Wenzhou 325000, China
[2]School of the 1st Clinical Medical Sciences, Wenzhou Medical University, Wenzhou 325035, China
[3]School of the 2nd Clinical Medical Sciences, Wenzhou Medical University, Wenzhou 325035, China
[4]College of Computer Science and Technology, Jilin University, Changchun 130012, China
[5]Department of Computer Science, Wenzhou University, Wenzhou 325035, China
[6]School of Digital Media, Shenzhen Institute of Information Technology, Shenzhen 518172, China
[7]School of Public Health and Management, Wenzhou Medical University, Wenzhou 325035, China

Corresponding authors: Huiling Chen (chenhuiling.jlu@gmail.com), Xuehua Zhao (lcrlc@sina.com), and Shu-Yun Zhang (shuyunzh@wmu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61602206, Grant 61471133, and Grant 61871475, in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY17F020012, in part by the Science and Technology Plan Project of Wenzhou, China, under Grant ZG2017019, in part by the Science and Technology Department of Zhejiang Province of China under Grant 2017C33128, in part by the Guangdong Natural Science Foundation under Grant 2018A030313339, in part by the Characteristic Innovation Projects of Universities in Guangdong under Grant 2017GKTSCX063, in part by the Ministry of Education in China Youth Fund Project of Humanities and Social Sciences under Grant 17YJCZH261, and in part by the 13th Five-Year Plan Project of Philosophy and Social Sciences in Shenzhen under Grant SZ2018D017.

ABSTRACT Pyrene, composed of four fused benzene rings, is a polycyclic aromatic hydrocarbon (PAH) that has served as a model compound for evaluating the toxic effects of PAHs. In this paper, 114 male rats were dosed daily by oral gavage with either vehicle (corn oil) or pyrene (1500 mg/kg/day) for four days. A method based on the gray wolf optimization-enhanced machine learning approach was then developed to identify pyrene poisoning in rats using the indices from blood analysis. The results showed that there were significant differences in blood analysis indices between the control and the pyrene groups ($p < 0.05$). In terms of feature selection, the most important correlated indices were liver to body weight ratio, the absolute value of leukomonocyte, the percentage of monocyte, serum albumin, direct bilirubin, urea nitrogen, and uric acid. This method was shown to have an accuracy rate of 94.62% (ACC), 0.8988 Matthews correlation coefficients (MCCs), 91.71% sensitivity, and 98.33% specificity. Empirical analysis indicates that this method represents a new and accurate approach for detecting pyrene poisoning.

INDEX TERMS Pyrene, hepatotoxicity, renal dysfunction, gray wolf optimization, feature selection, fuzzy KNN.

I. INTRODUCTION

Polycyclic aromatic hydrocarbons (PAHs) are severe pollutants mainly caused by incomplete combustion of various organic substances such as coal, gasoline, diesel, tobacco and foods [1]. When ingested or inhaled, PAHs are metabolized into compounds which, by binding slightly to DNA, form what is known as large DNA adducts, thus leading to mutations and carcinogenesis [2]. Pyrene is a polycyclic aromatic hydrocarbon composed of four fused benzene rings, and because it occurs at considerably high concentrations in PAH mixtures, it can serve as a biomarker of PAH exposure in human and wildlife [3], [4]. Pyrene exposure has lethal and sublethal effects, depending on dose. For instance, it can reduce feeding rate, oxygen uptake, growth rate, development time and reproductive yield. In addition, even at low concentrations (0.05-5nm), pyrene can cause cardiac abnormalities in fish such as pericardial edema and cardiac circulation defects, and a concentration-response experiment showed that, when its concentrations were higher than 100 nm [5], it resulted in reduced mobility in juvenile barramundi.

Zhang *et al.* [6] evaluated the negative effects of pyrene on the liver using wild-type mice orally dosed with pyrene, and found that pyrene exposure increased relative liver weight, induced hepatocellular hypertrophy and elevated ALT levels. Moreover, pyrene was found to induce expression of IL-8 in human epithelial cell lines [7].

Pyrene-induced hepatotoxicity, which is likely to involve various kinds of different liver injuries, has adverse impacts on human health, with a pivotal role in metabolism, the liver is extremely fragile to xenobiotics. Numerous computational models have been proposed to predict dose-dependent hepatotoxicity [8]. Liu *et al.* [9] constructed supervised machine learning to predict hepatotoxicity by using EPA ToxCast *in vitro* bioactivity and chemical structure based on six machine learning algorithms: classification and regression trees (CART), support vector machines (SVM), linear discriminant analysis (LDA), Naive Bayes (NB), k-nearest neighbors (KNN) and an ensemble of these classifiers (ENSMB). CART, SVM and ENSMB classifiers performed best in forecasting hepatotoxicity. Stoichiometry software was used to optimize the set of descriptors, and SVM model training was conducted for the end points of liver toxicity, achieving high sensitivity (68%) and good specificity (95%) [10]. Moreover, the SVM model, with an accuracy rate of 94.90% [11], was also established by extracting candidate data of specific biomarkers for hepatotoxicity and non-hepatotoxicity drugs. Ding *et al.* [12] used a model for hepatotoxicity prediction, which was based on a coarse-grained parallel genetic algorithm (CGPGA) and SVM, and the accuracy of this model was 78.21%.

Although drug-induced liver damage is not common, it remains an important safety issue that can lead to patient deaths and drug development failures [13], [14]. Toxicogenomics and quantitative structure-activity relationship (QSAR) modeling are widely used predictive tools in toxicology [15], [16]. Low *et al.* [17] developed multiple traditional QSAR classification models which use an integrated set of chemical descriptors and some classification methods such as KNN, SVM, random forests (RF) and distance weighting identification. Liew *et al.* [18] used the set construction model of mixed features and mixed learning algorithms to predict effects on the liver, achieving an accuracy rate of 75.0%, 81.9% sensitivity and 64.6% specificity. Lu *et al.* [19] used an approach based on the Bayesian reasoning method for prediction, and the average accuracy, sensitivity and specificity of the model were 78.47%, 74.17% and 82.77%, respectively. Using weighted fingerprint features, the prediction model was trained and evaluated by RF and SVM algorithms, achieving accuracy rates of 73.8% and 72.6%, respectively [20]. However, there are few reports of pyrene poisoning identification in animals or humans.

Blood analysis is one of the most common and widely available diagnostic methods in hospitals. In this study, using blood analysis indexes, a new method, based on the improved gray wolf optimization (GWO) machine learning approach, was developed to identify pyrene poisoning in rats. To our knowledge, our study is the first to adopt the GWO enhanced fuzzy k-nearest neighbor technique to identify pyrene poisoning in rats. As shown by the simulation results, the proposed GWO-FKNN approach effectively identified the most important features that can discriminate between pyrene-exposed rats and normal control rats. Comparison of the proposed method with two other nature-inspired metaheuristic algorithms-based FKNN methods demonstrates that the proposed GWO-FKNN can achieve better prediction results.

To sum up, the contributions of this paper are as follows: (1) create a new perspective of modeling pyrene poisoning in rats using blood samples and it works automatically and effectively; and (2) the most relevant indices can be identified by feature selection that is carried out using an improved GWO-based method.

II. MATERIALS AND METHODS

A. CHEMICALS AND ANIMALS

Pyrene (99% purity) was produced by Sigma-Aldrich (St. Louis, Missouri, USA). 114 male SD rats weighing 180-220 grams were from the Laboratory Animal Center of Wenzhou Medical University. The rats were kept at 23 ± 3°C with a 12-h light/dark cycle (light: 07:00-19:00 h, dark: 19:00-07:00 h) in the laboratory, and were given proper diets and enough water. All animal experiments were carried out in accordance with the Institutional Animal Care guidelines and were approved ethically by the Administration Committee of Experimental Animals, Laboratory Animal Center, Wenzhou Medical University.

B. TREATMENT OF ANIMALS

114 male SD rats weighing 180-220 grams were divided randomly into two groups, with 57 rats in the control group and 57 rats in the experimental group. Experimental rats were given daily by oral gavage with pyrene (dissolved in corn oil, 1500 mg/kg/day) for 4 days, whereas control rats received an equal volume of vehicle (corn oil). All rats were sacrificed 24 h after the last administration. Then the blood samples were collected for blood routine test (BRT) and biochemical analysis including serum alanine aminotransferase (ALT). Moreover liver and kidney weight was measured, and liver or kidneys/body weight ratios (LWR/KWR) are represented as (liver or kidneys weight)/(body mass) $\times$ 100%. Histologic analysis of hematoxylin-eosin staining was performed by fast dissection, cutting and fixation of liver tissues in 10% neutral phosphate buffer formaldehyde solution with hematoxylin-eosin staining.

1) LIST OF THE FEATURES

As illustrated in Table 1, a total of 36 blood indices were used in this study.

2) DATA ANALYSIS

Statistical analysis was performed using SPSS 17 software. The blood and biochemical indexes of the two groups were

TABLE 1. List of the features used in this study and their definitions.

Number	Features	Abbreviation
F1	liver to body weight ratio	LWR
F2	kidneys to body weight ratio	KWR
F3	white blood cell	WBC
F4	absolute value of leukomonocyte	VLC
F5	absolute value of neutrophile granulocyte	VNG
F6	absolute value of monocyte	VMC
F7	absolute value of eosinophils	VEO
F8	percentage of leukomonocyte	PLC
F9	percentage of neutrophile granulocyte	PNG
F10	percentage of monocyte	PMC
F11	percentage of eosinophils	PEO
F12	red blood cell	RBC
F13	hemoglobin	HGB
F14	hematokrit	HCT
F15	mean corpuscular volume	MCV
F16	red blood cell distribution width-standard deviation	SD
F17	red blood cell volume distribution width	RDW
F18	mean corpuscular hemoglobin	MCH
F19	mean corpuscular hemoglobin concentration	MCHC
F20	blood platelet	PLT
F21	mean platelet volume	MPV
F22	platelet distribution width	PDW
F23	platelet-large cell ratio	P-LCR
F24	thrombocytocrit	PCT
F25	serum alanine aminotransferase	ALT
F26	serum aspartate aminotransferase	AST
F27	AST to ALT ratio	AST/ALT
F28	alkaline phosphatase	ALP
F29	total serum protein	TP
F30	serum albumin	ALB
F31	serum globulin	GLB
F32	albumin to globulin ratio	A/G
F33	total bilirubin	TBIL
F34	direct bilirubin	DBIL
F35	indirect bilirubin	IBIL
F36	serum creatinine	SCR
F37	blood urea nitrogen	BUN
F38	blood uric acid	BUA

analyzed by a One-Way ANOVA test to discover statistical differences. P values less than 0.05 were considered statistically significant in all analyses.

C. METHODS

Figure 1 shows the flowchart of the method presented in this paper. The method consists of two parts: feature selection and classification. The blood samples were obtained from 57 pyrene exposed rats and 57 normal control rats. Each sample has 38 attributes, including 36 blood indices and 2 organ coefficients. Firstly, the data was normalized, and then the key features were selected from the data samples by GWO-FKNN method to optimize the parameters of the FKNN classifier, and then the optimal FKNN classifier was trained by using the optimal parameters and features obtained. Finally, the best

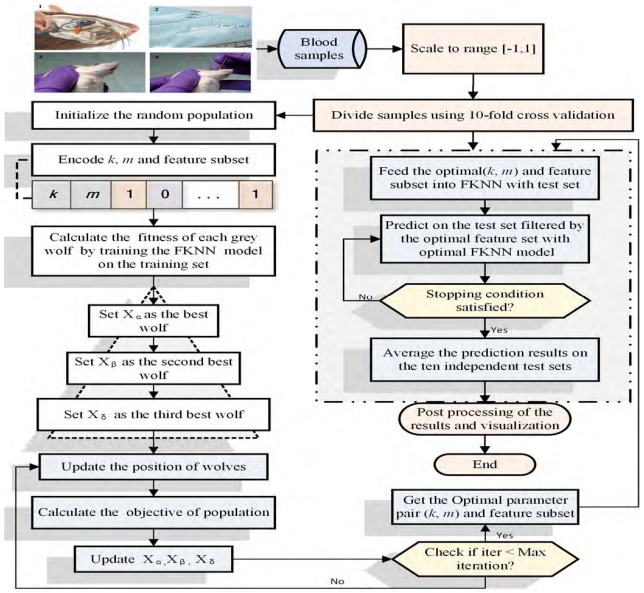


FIGURE 1. Flowchart of the proposed method.

FKNN classifier was used to predict the unknown samples. Data partition is based on 10-fold cross validation (CV).

1) PARAMETER OPTIMIZATION AND FEATURE SELECTION BY GRAY WOLF OPTIMIZATION (GWO)

In this study, a hybrid GWO-FKNN was constructed by combining the continuous GWO and the binary version of GWO, with the continuous GWO tackling the problem of parameter optimization and the binary GWO executing the feature selection task. GWO was first developed by Mirjalili *et al.* [21] as a new population-based heuristic search algorithm. Each swarm intelligence algorithm has its own characteristic [22]–[28], and GWO also has its unique characteristic in that it simulates the social rank and predatory behavior of gray wolves in nature [29].

The original GWO is known to be conducted in continuous space to solve the problems with real-valued parameters. However, the discrete binary version of GWO [30], [31] is commonly utilized to solve many optimization problems in engineering practice. Feature selection is a good instance of binary combinatorial optimization problems. In this study, the optimization task consists of binary variables and real-valued parameters, both of which need to be optimized. Therefore, we developed the hybrid GWO by combining the continuous GWO and the binary version of GWO in order that both real-valued and binary parameters of the task can be optimized.

2) CLASSIFICATION BASED ON FKNN

FKNN is a kernel predictive engine for executing the classification task after the optimal parameters and features are obtained. Compared with other traditional machine learning methods such as ANN and SVM, FKNN is much simpler

and can yield results that can be more easily interpreted. FKNN [32], [33] classifiers, which are based on conventional k-nearest neighbor (KNN) classifiers, have been studied extensively since first proposed.

For the past few years, FKNN has become a distinctive facet of neighbor classification and instance based learning [34]. Representing inaccurate information while providing the degree to the sample belonging to the relevant category is one of the most prominent features of FKNN. After being assigned a membership, each sample is classified to a class, which has the highest membership value. Owing to its fantastic traits, FKNN has been utilized in a broad array of scenarios, including medical diagnosis problems [35], [36], protein identification and prediction problems [37], [38], slope collapse prediction problems [39], bankruptcy prediction problems [40], and grouting activity prediction problems [41]. Nevertheless, FKNN has weaknesses in a certain sense, one of which is that it depends basically on two crucial parameters, the fuzzy intensity coefficient (m) and the number of neighbors (k). Hence, these two parameters have to be tuned appropriately in order to obtain better classification performance. In this study, we used GWO to tune the parameters of FKNN.

3) PROPOSED GWO-FKNN

To fully explore the potential of FKNN, GWO was used not only to optimize the parameters of FKNN, but also to identify the optimal feature subset in the data. The main steps of feature selection and parameter optimization via GWO-FKNN method are as follows:

Step 1: Initialize the input parameters for GWO, including maximum number of iterations, upper bounds of the variables, population size, and the dimension of the problem, lower bounds of the variables.

Step 2: Initialize a population of gray wolves randomly based on the upper and lower bounds of the variables.

Step 3: Initialize $a, \vec{A}$ and $\vec{C}$ using the following equations.

$$a = 2 \tag{1}$$

$$\vec{A} = 2a \cdot \vec{r}_1 - a \tag{2}$$

$$\vec{C} = 2\vec{r}_2 \tag{3}$$

where $\vec{r}_1$, $\vec{r}_2$ are random vectors in [0, 1].

Step 4: Encode each gray wolf with n+2 dimensions. The first two dimensions are parameter pair (k, m). The remaining n dimensions are discretized to binary values by equation 4, in which '1' indicates the feature is selected, and '0' indicates the feature is not selected.

$$flag_{i,j} = \begin{cases} 1 & X_{i,j} > 0.5 \\ 0 & otherwise \end{cases} \tag{4}$$

where $X_{i,j}$ indicates the jth position of the ith gray wolf.

Step 5: Calculate the fitness with (k, m) and the selected features for each gray wolf according to the following equation.

$$\begin{cases} f_1 = \dfrac{\sum_{i=1}^{K} acc_i}{K} \\ f_2 = 1 - \dfrac{\sum_{j=1}^{n} bin_j}{n} \\ f = \alpha \times f_1 + \beta \times f_2 \end{cases} \tag{5}$$

The first sub-objective function f_1 represents the average accuracy achieved by the FKNN classifier via K-fold CV, where $K = 5$ and acc_i is the accuracy of the ith fold CV. In the second sub-objective function f_2, bin_j is the binary value of the jth feature, n is the total number of features. In objective function f, α is the weight for FKNN classification accuracy, β indicates the weight for the selected features.

Step 6: Choose the first three best gray wolves that have maximum fitness and save them as α, β, and δ.

Step 7: Update the position of the rest of the population (ω) using equations 6 to 12.

$$\vec{D}_\alpha = \left| \vec{C}_1 \cdot \vec{X}_\alpha - \vec{X} \right| \tag{6}$$

$$\vec{D}_\beta = \left| \vec{C}_2 \cdot \vec{X}_\beta - \vec{X} \right| \tag{7}$$

$$\vec{D}_\delta = \left| \vec{C}_3 \cdot \vec{X}_\delta - \vec{X} \right| \tag{8}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha \tag{9}$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta \tag{10}$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot \vec{D}_\delta \tag{11}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{12}$$

where t indicates the current iteration.

Step 8: Return back if the gray wolves go beyond the boundaries of the variables.

Step 9: Decrease the value of a from 2 to 0 linearly.

Step 10: Update $\vec{A}$ and $\vec{C}$ using the equation 2 and 3 respectively.

Step 11: Go to step 4 if the end criterion is not satisfied.

Step 12: Return the first two dimensions' continuous values of X_α as the optimal FKNN parameter pair (k, m) and the remaining n dimensions' binary values of X_α as the markers of the best feature subset.

III. EXPERIMENTAL SETUP

The experiment was implemented in MATLAB on a Windows Server 2008 R2 operating system with Intel (R) Xeon (R) CPU E5-2660 v3 (2.60 GHz) and 16GB of RAM.

GA, GWO, FKNN and PSO were implemented from the beginning.

Data were firstly normalized into the range $[-1, 1]$ prior to building the prediction model. We used the stratified 10-fold CV [42] to evaluate classification performance in order to guarantee unbiased results. The number of the maximum iterations and search agent size for GWO, PSO, GA were fixed at 50 and 20. The two constant factors c_1 and c_2 in PSO were set to 2, the inertial weight w in PSO was set to 1. Crossover fraction and mutation probability in GA were set to 0.8 and 0.01, respectively. The searching ranges for the two parameters in FKNN were set as follows: $k \in [1, 5]$, $m \in [1, 5]$.

Evaluation criteria including specificity, sensitivity, classification accuracy (ACC) and Matthews Correlation Coefficients (MCC) were analyzed in order to evaluate the proposed method. They are defined as follows:

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \quad (13)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (14)$$

$$Specificity = \frac{TN}{FP + TN} \times 100\% \quad (15)$$

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \times 100\% \quad (16)$$

IV. EXPERIMENTAL RESULTS AND CONCLUSIONS

A. EXPERIMENT I: CLASSIFICATION WITHOUT FEATURE SELECTION

Figure 2 shows the changes in tissue structure and collagen. There were severe degenerations in the hepatocytes of pyrene-treated rats such as cellular swelling. However, in the control rats, only normal liver cells, central veins and a small amount of collagen fibers were observed.

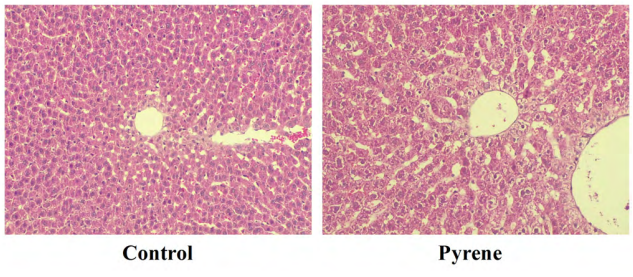


FIGURE 2. Pyrene exposure causes pathological changes to rat liver tissue (HE staining × 200). Two groups rats were orally administered with corn oil (control) or pyrene (1500 mg/kg) once daily for 4 days.

Table 2 lists the blood indices in both the control and pyrene-treated groups. The activity of serum ALT notably increased in rats exposed to pyrene ($P < 0.05$), suggesting pyrene exposure induced hepatocyte injury.

TABLE 2. Statistical analysis of 114 rats.

Indexes		Control(n=57)	Pyrene (n=57)	p-Value
LWR	%	3.92±0.55	5.06±0.51	0.000
KWR	%	0.81±0.09	0.87±0.09	0.001
WBC	10^9/L	7.60±1.92	9.72±3.23	0.000
VLC	10^9/L	6.06±1.89	7.61±2.71	0.001
VNG	10^9/L	1.15±0.81	1.42±1.15	0.142
VMC	10^9/L	0.29±0.19	0.59±0.42	0.000
VEO	10^9/L	0.10±0.06	0.09±0.05	0.086
PLC	%	79.44±10.89	78.42±10.39	0.611
PNG	%	15.42±10.28	14.63±9.08	0.663
PMC	%	3.69±1.98	6.00±3.19	0.000
PEO	%	1.45±1.04	0.94±0.54	0.001
RBC	10^{12}/L	7.93±0.87	7.96±0.66	0.790
HGB	g/L	148.12±11.81	153.23±10.68	0.017
HCT	%	44.46±3.17	45.33±2.70	0.116
MCV	fL	56.46±4.21	57.09±2.94	0.361
SD	fL	29.00±2.04	29.46±2.06	0.233
RDW	%	16.80±2.09	16.84±1.47	0.897
MCH	pg	18.77±0.94	19.28±0.73	0.002
MCHC	g/L	333.12±12.24	337.89±8.04	0.016
PLT	10^9/L	1070.26±154.02	1233.89±200.09	0.000
MPV	fL	7.00±0.20	6.94±0.22	0.108
PDW	fL	7.52±0.32	7.46±0.44	0.426
P-LCR	%	5.82±1.21	5.62±1.24	0.369
PCT	%	0.75±0.10	0.86±0.14	0.000
ALT	U/L	35.47±8.96	51.77±20.31	0.000
AST	U/L	101.70±32.26	90.65±24.59	0.042
AST/ALT	/	3.12±1.59	1.94±0.81	0.000
ALP	U/L	179.65±53.82	217.68±77.25	0.003
TP	g/L	56.12±3.49	57.95±4.26	0.014
ALB	g/L	33.11±2.78	33.39±2.78	0.591
GLB	g/L	23.02±2.10	24.56±3.17	0.003
A/G	/	1.45±0.17	1.37±0.16	0.013
TBIL	μmol/L	1.73±0.22	1.91±0.42	0.006
DBIL	μmol/L	1.49±0.25	1.83±0.44	0.000
IBIL	μmol/L	0.24±0.27	0.07±0.10	0.000
SCR	μmol/L	39.33±6.69	45.32±12.28	0.002
BUN	mmol/L	8.44±1.59	12.38±3.82	0.000
BUA	μmol/L	110.18±39.27	96.05±31.41	0.036

LWR is known to effectively reflect liver pathological changes and nutritional status in the body. Table 2 shows the LWR and KWR in the two groups. Significant increases in LWR and KWR were observed in pyrene-treated rats, compared to the control group ($P < 0.05$). Therefore, the results suggested that pyrene exposure results in hepatomegaly and renal enlargement in rats.

Table 3 shows the classification results obtained by the GWO-FKNN method. As can be seen from the table, GWO-FKNN yields 94.62% ACC, 91.71% sensitivity, 98.33% specificity and 0.8988 MCC, indicating that the method, by making full use of GWO's powerful search and optimization ability to identify the first-rank combination of parameters and features in training data, can set up the optimal forecasting model to predict the new samples.

To illustrate the effect of feature selection, Figure 3 lists the comparison results between GWO-FKNN with and without feature selection. As can be seen from the figure compared

TABLE 3. The detailed results obtained by GWO-FKNN.

Fold	Selected feature subset	ACC	MCC	Sensitivity	Specificity
#1	{1,3,5,8,9,12,16,17,19,20,23, 24,25,27,30,32,33,34,35,36, 37,38}	1.0000	1.0000	1.0000	1.0000
#2	{1,4,5,10,11,15,20,23,24,25, 28,30,31,32,33,34,35,37,38 }	1.0000	1.0000	1.0000	1.0000
#3	{1,3,4,6,7,8,10,12,14,15,19, 21,24,25,26,27,28,31,33,35, 37}	1.0000	1.0000	1.0000	1.0000
#4	{1,2,4,5,7,9,10,11,16,20,21, 22,28,29,30,32,34,37,38}	0.9091	0.8571	1.0000	0.8281
#5	{1,2,4,6,7,8,10,15,17,19,20, 21,23,24,27,30,31,33,36,37 }	0.8182	0.7143	1.0000	0.6901
#6	{1,4,5,9,10,15,21,22,25,28,3 0,31,32,33,34,36,37,38}	1.0000	1.0000	1.0000	1.0000
#7	{1,2,4,5,10,15,21,24,25,27,2 9,30,34,37,38}	0.8182	0.8000	0.8333	0.6333
#8	{1,4,5,10,11,12,19,20,22,25, 28,29,30,32,33,34,37,38}	1.0000	1.0000	1.0000	1.0000
#9	{1,2,3,8,9,11,12,15,16,18,19, 20,21,23,27,30,32,34,36,37, 38}	1.0000	1.0000	1.0000	1.0000
#10	{1,2,3,4,6,10,11,16,19,20,21, 22,25,28,29,30,34,35,36,37, 38}	0.9167	0.8000	1.0000	0.8367
Avg.	-	0.9462	0.9171	0.9833	0.8988
Std.	-	0.0763	0.1122	0.0527	0.1431

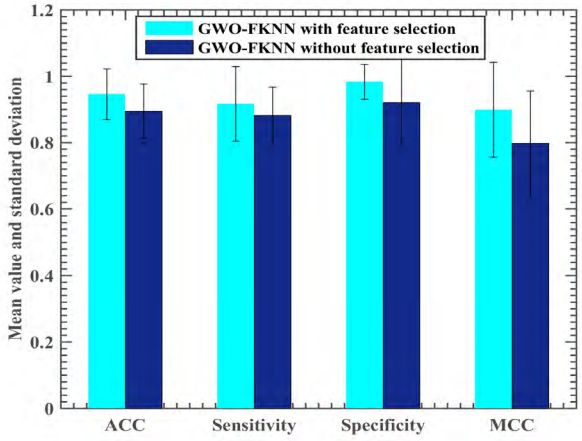


FIGURE 3. Comparison of GWO-FKNN with feature selection and without feature selection.

with the GWO-FKNN method without feature selection, ACC, sensitivity, specificity and MCC are improved with feature selection by 5.07%, 3.48%, 6.16% and 10.09%, respectively, indicating that feature selection leads to an overall performance improvement.

In addition, the standard deviation of the model with a feature selection function on ACC, sensitivity and specificity is much smaller than the other model, indicating that feature selection can enhance the stability of the model to a certain extent. After feature selection, we found that 7 features including LWR, VLC, PMC, ALB, DBIL, BUN and BUA are the most frequent noes appearing in the 10 feature subsets in Figure 4. This shows that these features play a key role in distinguishing poisoned rats from normal rats, and that there

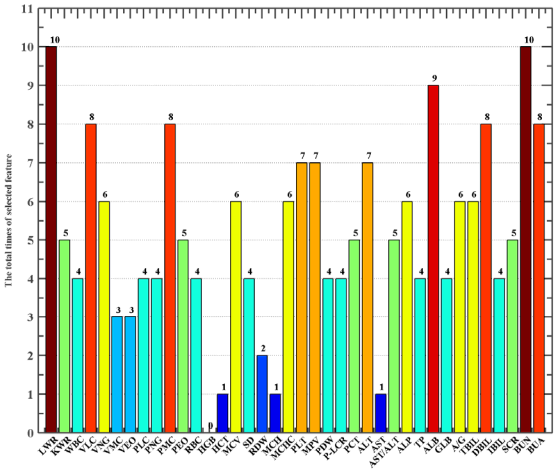


FIGURE 4. The frequencies of selected features during 10-fold CV procedure.

are some redundant and unrelated features in the experimental data.

To further demonstrate the effectiveness of the proposed GWO-FKNN method, GWO-FKNN was compared with PSO-FKNN and GA-FKNN. Figure 5 shows the performance of these three approaches in terms of the four indicators. As can be seen from this figure, GWO-FKNN outperforms the other two competitors in terms of the four indicators. Moreover, the standard deviation of GWO-FKNN is smaller than that of the other two methods. These data show that the proposed GWO-FKNN is a promising method for distinguishing poisoned rats from normal rats.

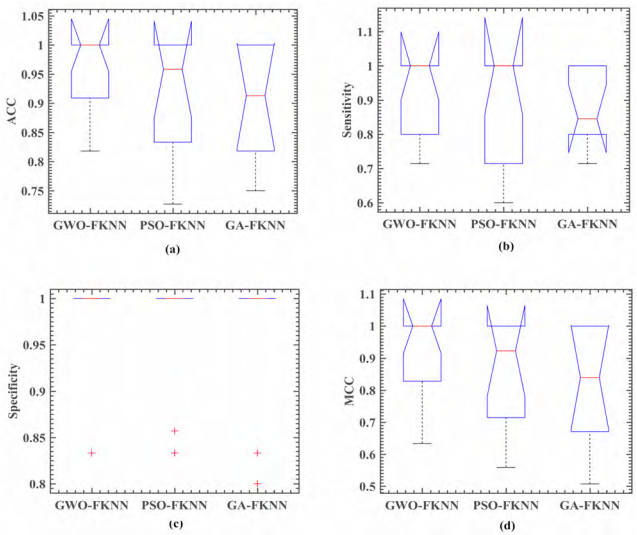


FIGURE 5. Boxplot of classification performance obtained by the four methods in terms of ACC(a), sensitivity(b), specificity(c) and MCC(d).

Figure 6 depicts the convergence trends of GWO-FKNN, PSO-FKNN and GA-FKNN. As can be seen from the figure, GWO-FKNN has the fastest convergence speed and obtains the best solution among the three methods. Notably, the best

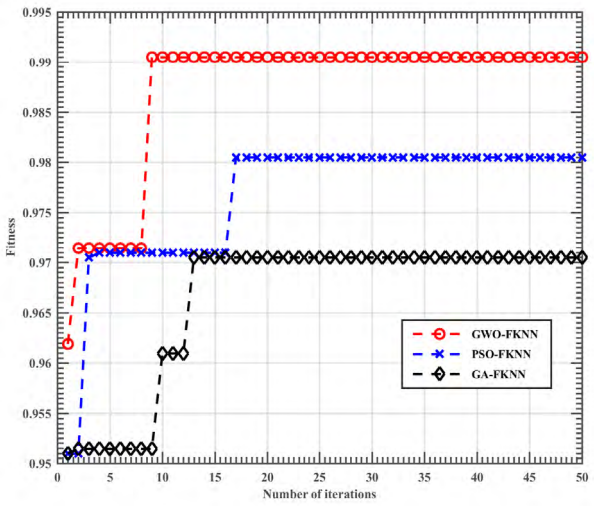


FIGURE 6. The convergent evolution trend chart of the three methods.

solution is found by GWO-FKNN in less than 10 iterations. This may be attributed mainly to GWO's better global search and local exploration ability in searching for optimal parameter and feature combinations. Compared with GWO-FKNN, PSO-FKNN has a relatively slow convergence speed, and it takes more than 15 iterations to find the best solution, which is smaller than that of GWO-FKNN. Although GA-FKNN converges faster than PSO-FKNN, it converges to the suboptimal solution, which may result from the entrapment in local optimum during the search process.

V. DISCUSSION AND FUTURE WORK

PAHs have been identified as endocrine disrupters, reproductive toxicants, cardiovascular toxicants, neural toxicants and carcinogens [43]–[45]. To the best of our knowledge, our study is the first to apply the GWO-FKNN method to identify pyrene poisoning in rats based on blood, liver and kidney indices. Analysis of variance suggested that WBC, LWR, VLC, KWR, HGB, PMC, MCH, MCHC, ALP, VMC, PEO, PLT, ALT, AST, PCT, TP, AST/ALT, SCR, IBIL, GLB, A/G, DBIL, TBIL, BUN and BUA all identify the differences between the control and pyrene groups. The p values of all indices (excepting HGB, MCHC, AST, TP, A/G and BUA) are less than 0.01, indicating these indices are likely to have much more important toxicological significance. However, it is difficult to determine the most important indices, such as accurate classification of gliomas is crucial for prescribing therapy and assessing the prognosis of patients. Wang *et al.* [46] developed and validated that multiparametric MRI-based radiomic analysis is a novel and convenient approach for the classification of gliomas into low-grade and high-grade tumors. Tian *et al.* [47] pointed out that radiomic strategy incorporating 3D texture features from multiparametric MRI was highly effective for noninvasive glioma grading. As can be seen from Table 3, GWO-FKNN approach achieved the highest classification accuracy (94.62%) when the feature set included several frequent features in the process of feature selection. The feature selection results show that VLC, PMC, LWR, BUN, DBIL, ALB and BUA are more important than WBC, VMC, KWR, PLT, IBIL, ALT, MCH, PEO, TBIL, PCT, SCR, AST/ALT, ALP, GLB and AST.

In this study, we found that pyrene poisoning increases LWR and results in notable hepatic histological changes in rats. Previous studies have suggested that pyrene exposure prominently increases the relative liver weight and causes tissue injuries such as hepatocellular hypertrophy in mice [6]. Likewise, we observed that oral administration of pyrene at a dose of 1500 mg/kg markedly increases liver index in rats.

It is common to use hematologic and biochemical analyses to assess organ dysfunction in toxicological safety assessments [48]. It has been proved that some hematologic parameters reflect inflammatory responses. For instance, leukomonocyte and monocyte have been shown to be related to undesirable outcomes in some solid tumors such as ovarian, colorectal, gastric, cervical, breast and oesophagal carcinomas [49]. Leukomonocytes, including B cells, T cells and NK cells, are one of the most vital immune cells whose main duty is to identify and kill viruses, tumor cells and bacteria. When attacked by viruses, leukomonocytes can soon be activated and cause a chain of immune reactions [50]. Zhu *et al.* [51] found leukomonocyte count in rats whose myocardial infarction was acute was prominently higher than that in the normal control group ($p < 0.05$). The most vital cells of the innate immune system in the peripheral blood are monocytes. Gestational diabetes Mellitus (GDM) is known to be an inflammatory disease which involves different kinds of cells and mediators during its development. Angelo *et al.* [52] demonstrated that the profile of monocytes is altered in the peripheral blood of GDM patients. Machida *et al.* [53] showed that increased monocyte count is an independent predictor of tumor recurrence and progression.

Serum albumin (ALB), the most abundant protein in plasma, is synthesized in the liver. ALB plays a major role in detoxifying reactive oxygen species [54]. Serum bilirubin, with anti-inflammatory properties, is an end product of heme metabolism and an effective endogenous antioxidant. There are three forms of bilirubin in plasma: unconjugated bilirubin (usually the main ingredient; reversibly bound to albumin), free conjugated bilirubin and delta bilirubin (conjugated bilirubin, also known as direct bilirubin (DBIL), normal range approximately 0.0–3.0 μmol/L; covalently bound to albumin). Hodgson *et al.* [55] showed that, among clinically relevant population of disease-free neonates with prolonged jaundice, both the total and the direct bilirubin decrease with age. However, the absolute direct bilirubin is more clinically useful than the direct-total bilirubin ratio. Song *et al.* [56] showed DBIL levels are not only associated with tumor progression and first-line platinum chemotherapy reactions, but also are correlated with total survival after lymph node metastasis.

Blood urea nitrogen (BUN) was demonstrated to be helpful for risk stratification in a variety of diseases such as

acute myocardial infarction, heart failure and diabetes mellitus [57]–[59]. Given its reabsorption by the tubules and its role in physiological fluid balance [60], BUN is regarded as a marker for not only renal function but also neurohumoral activity. It was suggested that blood uric acid (BUA) is one of the causes for both kidney damage and renal function progression [61]. Since the kidney plays an important role in regulating BUA levels, hyperuricemia is primarily caused by increased intake or secretion defects resulting from excessive deposition of crystals in proximal renal tubules. However, hyperuricemia is also related to metabolic syndromes, diabetes, renal injuries, hypertension and cardiovascular diseases.

In this study, we found multiple useful indices that can be used for identifying pyrene poisoning, including VLC, ALB, LWR, DBIL, PMC, BUN and BUA. However, whether these indices can be used for identifying pyrene poisoning in humans remains to be determined. In order to offer higher accuracy of prediction, more data samples need to be collected. In addition, we are attempting to create an expert system that can automatically identify pyrene poisoning based on the GWO-FKNN method.

ACKNOWLEDGMENT

(Huiling Chen, Xuehua Zhao, and Shu-Yun Zhang contributed equally to this work.)

REFERENCES

[1] C. C. Tonne, R. M. Whyatt, D. E. Camann, F. P. Perera, and P. L. Kinney, "Predictors of personal polycyclic aromatic hydrocarbon exposures among pregnant minority women in New York City," *Environ. Health Perspect.*, vol. 112, no. 6, pp. 754–759, May 2004.

[2] P. Irigaray and D. Belpomme, "Basic properties and molecular mechanisms of exogenous chemical carcinogens," *Carcinogenesis*, vol. 31, no. 2, pp. 135–148, Feb. 2010.

[3] E. Dam, B. Styrishave, K. F. Rewitz, and O. Andersen, "Intermoult duration affects the susceptibility of shore crabs Carcinus maenas (L.) to pyrene and their ability to metabolise it," *Aquatic ToxiCology*, vol. 80, no. 3, pp. 290–297, Dec. 2006.

[4] C. Viau, M. Bouchard, G. Carrier, R. Brunet, and K. Krishnan, "The toxicokinetics of pyrene and its metabolites in rats," *ToxiCology Lett.*, vol. 108, nos. 2–3, pp. 201–207, Sep. 1999.

[5] O. Guven, L. Bach, P. Munk, K. V. Dinh, P. Mariani, and T. G. Nielsen, "Microplastic does not magnify the acute effect of PAH pyrene on predatory performance of a tropical fish (Lates calcarifer)," *Aquatic ToxiCology*, vol. 198, pp. 287–293, May 2018.

[6] X. J. Zhang, Z. Shi, J. X. Lyv, X. He, N. A. Englert, and S. Y. Zhang, "Pyrene is a novel constitutive androstane receptor (CAR) activator and causes hepatotoxicity by CAR," *Toxicological Sci.*, vol. 147, no. 2, pp. 436–445, Oct. 2015.

[7] H. Bömmel *et al.*, "The diesel exhaust component pyrene induces expression of IL-8 but not of eotaxin," *Int. Immunopharmacol.*, vol. 3, nos. 10–11, pp. 1371–1379, Oct. 2003.

[8] A. Cheng and S. L. Dixon, "*In silico* models for the prediction of dose-dependent human hepatotoxicity," *J. Comput.-Aided Mol. Des.*, vol. 17, no. 12, pp. 811–823, Dec. 2003.

[9] J. Liu *et al.*, "Predicting hepatotoxicity using ToxCast in vitro bioactivity and chemical structure," *Chem. Res. Toxicology*, vol. 28, no. 4, pp. 738–751, Feb. 2015.

[10] D. Mulliner, F. Schmidt, M. Stolte, H.-P. Spirkl, A. Czich, and A. Amberg, "Computational models for human and animal hepatotoxicity with a global application scope," *Chem. Res. Toxicology*, vol. 29, no. 5, pp. 757–767, Feb. 2016.

[11] Y. Li *et al.*, "A systematic strategy for screening and application of specific biomarkers in hepatotoxicity using metabolomics combined with ROC curves and SVMs," *Toxicological Sci.*, vol. 150, no. 2, pp. 390–399, Apr. 2016.

[12] S. Ding, S.-Y. Zhao, Z. Chen, and T. Lin, "A model for hepatotoxicity prediction based on coarse-grained parallel genetic algorithm and support vector machine," *J. Med. Imag. Health Inform.*, vol. 6, no. 8, pp. 1896–1903, Dec. 2016.

[13] M. Z. Sakatis *et al.*, "Preclinical strategy to reduce clinical hepatotoxicity using in vitro bioactivation data for >200 compounds," *Chem. Res. Toxicology*, vol. 25, no. 10, pp. 2067–2082, Oct. 2012.

[14] L. Ye *et al.*, "Predicting hepatotoxicity of compounds from traditional chinese medicines using tree models," *Chin. Pharmaceutical J.*, vol. 49, pp. 1583–1588, 2014.

[15] P. Zhao, B. Liu, and C. Wang, "Hepatotoxicity evaluation of traditional Chinese medicines using a computational molecular model," *Clin. Toxicology*, vol. 55, no. 9, pp. 996–1000, Nov. 2017.

[16] X. Zhu and N. L. Kruhlak, "Construction and analysis of a human hepatotoxicity database suitable for QSAR modeling using post-market safety data," *Toxicology*, vol. 321, pp. 62–72, Jul. 2014.

[17] Y. Low *et al.*, "Predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches," *Chem. Res. Toxicology*, vol. 24, no. 8, pp. 1251–1262, Jul. 2011.

[18] C. Y. Liew, Y. C. Lim, and C. W. Yap, "Mixed learning algorithms and features ensemble in hepatotoxicity prediction," *J. Comput.-Aided Mol. Des.*, vol. 25, pp. 855–871, Sep. 2011.

[19] Y. Lu *et al.*, "Predicting hepatotoxicity of drug metabolites via an ensemble approach based on support vector machine," *Combinat. Chem. High Throughput Screening*, vol. 20, no. 10, pp. 839–849, Dec. 2017.

[20] E. Kim and H. Nam, "Prediction models for drug-induced hepatotoxicity by using weighted molecular fingerprints," *BMC Bioinf.*, vol. 18, no. 7, p. 227, May 2017.

[21] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.

[22] J. Luo, H. Chen, Q. Zhang, Y. Xu, H. Huang, and X. Zhao, "An improved grasshopper optimization algorithm with application to financial stress prediction," *Appl. Math. Model.*, vol. 64, pp. 654–668, Dec. 2018.

[23] Z. Cai *et al.*, "An intelligent parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach," *Comput. Math. Methods Med.*, vol. 2018, Jun. 2018, Art. no. 2396952.

[24] W. Deng, H. Zhao, L. Zou, G. Li, X. Yang, and D. Wu, "A novel collaborative optimization algorithm in solving complex optimization problems," *Soft Comput.*, vol. 21, no. 15, pp. 4387–4398, Aug. 2017.

[25] W. Deng, H. Zhao, X. Yang, J. Xiong, M. Sun, and B. Li, "Study on an improved adaptive PSO algorithm for solving multi-objective gate assignment," *Appl. Soft Comput.*, vol. 59, pp. 288–302, Oct. 2017.

[26] W. Deng, R. Yao, H. Zhao, X. Yang, and G. Li, "A novel intelligent diagnosis method using optimal LS-SVM with improved PSO algorithm," *Soft Comput.*, pp. 1–18, 2017.

[27] Q. Zhang, H. Chen, J. Luo, Y. Xu, C. Wu, and C. Li, "Chaos enhanced bacterial foraging optimization for global optimization," *IEEE Access*, vol. 6, pp. 64905–64919, 2018.

[28] X. Wang, Z. Wang, J. Weng, C. Wen, H. Chen, and X. Wang, "A new effective machine learning framework for sepsis diagnosis," *IEEE Access*, vol. 6, pp. 48300–48310, 2018, doi: 10.1109/ACCESS.2018.2867728.

[29] X. Zhao *et al.*, "Chaos enhanced grey wolf optimization wrapped ELM for diagnosis of paraquat-poisoned patients," *Comput. Biol. Chem.*, to be published.

[30] Q. Li *et al.*, "An enhanced grey wolf optimization based feature selection wrapped kernel extreme learning machine for medical diagnosis," *Comput. Math. Methods Med.*, vol. 2017, Jan. 2017, Art. no. 9512741.

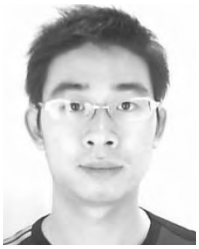
[31] Y. Wei *et al.*, "An improved grey wolf optimization strategy enhanced SVM and its application in predicting the second major," *Math. Problems Eng.*, vol. 2017, Feb. 2017, Art. no. 9316713.

[32] A. Jóźwik, "A learning scheme for a fuzzy k-NN rule," *Pattern Recognit. Lett.*, vol. 1, nos. 5–6, pp. 287–289, Jul. 1983.

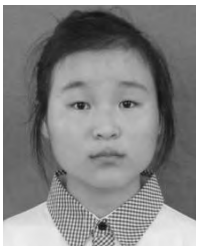
[33] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy K-nearest neighbor algorithm," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 4, pp. 580–585, Jul./Aug. 1985.

[34] J. Derrac, S. García, and F. Herrera, "Fuzzy nearest neighbor algorithms: Taxonomy, experimental analysis and prospects," *Inf. Sci.*, vol. 260, pp. 98–119, Mar. 2014.

[35] H.-L. Chen *et al.*, "An efficient diagnosis system for detection of Parkinson's disease using fuzzy *k*-nearest neighbor approach," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 263–271, Jan. 2013.
[36] D. Y. Liu, H. L. Chen, B. Yang, X. E. Lv, L. N. Li, and J. Liu, "Design of an enhanced fuzzy k-nearest neighbor classifier based computer aided diagnostic system for thyroid disease," *J. Med. Syst.*, vol. 36, no. 5, pp. 3243–3254, Oct. 2012.
[37] Y. Huang and Y. Li, "Prediction of protein subcellular locations using fuzzy k-NN method," *Bioinformatics*, vol. 20, no. 1, pp. 21–28, Jan. 2004.
[38] J. Sim, S. Y. Kim, and J. Lee, "Prediction of protein solvent accessibility using fuzzy k-nearest neighbor method," *Bioinformatics*, vol. 21, no. 12, pp. 2844–2849, Jul. 2005.
[39] M.-Y. Cheng and N.-D. Hoang, "A swarm-optimized fuzzy instance-based learning approach for predicting slope collapses in mountain roads," *Knowl.-Based Syst.*, vol. 76, pp. 256–263, Mar. 2015.
[40] H. L. Chen *et al.*, "A novel bankruptcy prediction model based on an adaptive fuzzy *k*-nearest neighbor method," *Knowl.-Based Syst.*, vol. 24, no. 8, pp. 1348–1359, Dec. 2011.
[41] M.-Y. Cheng and H. Nhat-Duc, "Groutability estimation of grouting processes with microfine cements using an evolutionary instance-based learning approach," *J. Comput. Civil Eng.*, vol. 28, no. 4, p. 04014014, Jul. 2014.
[42] S. L. Salzberg, "On comparing classifiers: Pitfalls to avoid and a recommended approach," *Data Mining Knowl. Discovery*, vol. 1, no. 3, pp. 317–328, Sep. 1997.
[43] C. L. Davie-Martin, K. G. Stratton, J. G. Teeguarden, K. M. Waters, and S. L. M. Simonich, "Implications of bioremediation of polycyclic aromatic hydrocarbon-contaminated soils for human health and cancer risk," *Environ. Sci. Technol.*, vol. 51, no. 17, pp. 9458–9468, Aug. 2017.
[44] J. Korsh, A. Shen, K. Aliano, and T. Davenport, "Polycyclic aromatic hydrocarbons and breast cancer: A review of the literature," *Breast Care*, vol. 10, no. 5, pp. 316–318, Oct. 2015.
[45] P. Poursafa *et al.*, "A Systematic review on the effects of polycyclic aromatic hydrocarbons on cardiometabolic impairment," *Int. J. Prev. Med.*, vol. 8, p. 19, Apr. 2017.
[46] Q. Wang *et al.*, "Radiomics nomogram building from multiparametric mrito predict grade in patients with glioma: A cohort study," *J. Magn. Reson. Imag.*, p. 8, Jun. 2018.
[47] Q. Tian *et al.*, "Radiomics strategy for glioma grading using texture features from multiparametric MRI," *J. Magn. Reson. Imag.*, vol. 48, no. 6, pp. 1518–1528, Dec. 2018.
[48] Q. He *et al.*, "Sex-specific reference intervals of hematologic and biochemical analytes in *Sprague-Dawley* rats using the nonparametric rank percentile method," *PLoS ONE*, vol. 12, p. e0189837, Dec. 2017.
[49] I. Temur, U. Kucukgoz Gulec, S. Paydas, A. B. Guzel, M. Sucu, and M. A. Vardar, "Prognostic value of pre-operative neutrophil/lymphocyte ratio, monocyte count, mean platelet volume, and platelet/lymphocyte ratio in endometrial cancer," *J. Obstetrics Gynecology Reproductive Biol.*, vol. 226, pp. 25–29, Jul. 2018.
[50] Y. Song, Y. Shen, X. Xia, and A.-M. Zhang, "Association between genetic polymorphisms of the IL28B gene and leukomonocyte in chinese hepatitis B virus-infected individuals," *PeerJ*, vol. 5, p. e4149, Dec. 2017.
[51] G. Zhu, Y. Yao, L. Pan, W. Zhu, and S. Yan, "Reduction of leukocyte counts by hydroxyurea improves cardiac function in rats with acute myocardial infarction," *Med. Sci. Monitor*, vol. 21, pp. 3941–3947, Dec. 2015.
[52] A. G. S. Angelo *et al.*, "Monocyte profile in peripheral blood of gestational diabetes mellitus patients," *Cytokine*, vol. 107, pp. 79–84, Jul. 2018.
[53] H. Machida, M. Y. De Zoysa, T. Takiuchi, M. S. Hom, K. E. Tierney, and K. Matsuo, "Significance of monocyte counts at recurrence on survival outcome of women with endometrial cancer," *Int. J. Gynecological Cancer*, vol. 27, no. 2, pp. 302–310, Feb. 2017.
[54] F. Li, M. D. Chordia, K. A. Woodling, and T. L. Macdonald, "Irreversible alkylation of human serum albumin by zileuton metabolite 2-acetylbenzothiophene-S-oxide: A potential model for hepatotoxicity," *Chem. Res. Toxicol*, vol. 20, no. 12, pp. 1854–1861, Dec. 2007.
[55] J. M. Hodgson, V. H. van Someren, C. Smith, and A. Goyale, "Direct bilirubin levels observed in prolonged neonatal jaundice: A retrospective cohort study," *BMJ Paediatrics Open*, vol. 2, no. 1, p. e000202, Feb. 2018.
[56] Y.-J. Song, X.-H. Gao, Y.-Q. Hong, and L.-X. Wang, "Direct bilirubin levels are prognostic in non-small cell lung cancer," *Oncotarget*, vol. 9, no. 1, pp. 892–900, Jan. 2018.
[57] O. Arihan *et al.*, "Blood Urea nitrogen (BUN) is independently associated with mortality in critically ill patients admitted to ICU," *PLoS ONE*, vol. 13, no. 1, p. e0191697, Jan. 2018.
[58] Y. Horiuchi *et al.*, "A high level of blood urea nitrogen is a significant predictor for in-hospital mortality in patients with acute myocardial infarction," *Int. Heart J.*, vol. 59, no. 2, pp. 263–271, Mar. 2018.
[59] Y. Xie, B. Bowe, T. Li, H. Xian, Y. Yan, and Z. Al-Aly, "Higher blood urea nitrogen is associated with increased risk of incident diabetes mellitus," *Kidney Int.*, vol. 93, no. 3, pp. 741–752, Mar. 2018.
[60] B. Wernly *et al.*, "Blood urea nitrogen (BUN) independently predicts mortality in critically ill patients admitted to ICU: A multicenter study," *Clin. Hemorheology Microcirculation*, vol. 69, nos. 1–2, pp. 123–131, Jan. 2018.
[61] H. Y. Lu *et al.*, "Predictive Value of Serum Creatinine, Blood Urea Nitrogen, Uric Acid, and β2-microglobulin in the evaluation of acute kidney injury after orthotopic liver transplantation," *Chin. Med. J.*, vol. 131, no. 9, pp. 1059–1066, May 2018.



JIAYIN ZHU is currently a Lecturer with the Laboratory Animal Center, Wenzhou Medical University, China. His main research interests include laboratory animal and pesticide toxicology.

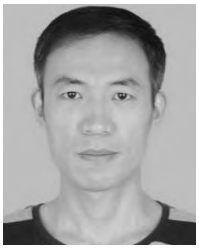


FENGTING ZHU is currently pursuing the degree with the School of the 1st Clinical Medical Sciences, Wenzhou Medical University. Her main research interest includes medical diagnosis.

SIYI HUANG is currently pursuing the degree with the School of the 2nd Clinical Medical Sciences, Wenzhou Medical University. Her main research interest includes medical diagnosis.

GANG WANG received the Ph.D. degree in theoretical computer science from Jilin University. He joined the College of Geo-Exploration Science and Technology Moving Platform Exploration Technology Research and Development Center, Jilin University, as a Postdoctoral Researcher. He is currently an Associate Professor with the College of Computer Science and Technology, Jilin University. His research interests include computer vision and machine learning.

HUILING CHEN received the Ph.D. degree from the Department of Computer Science and Technology, Jilin University, China. He is currently an Associate Professor with the Department of Computer Science and Technology, Wenzhou University, China. He has published more than 100 papers in international journals and conference proceedings, including *Pattern Recognition*, *Expert Systems With Applications*, *Knowledge-Based Systems*, *Soft Computing*, *Neurocomputing*, PAKDD, and among others. His present research interests include machine learning and data mining and their applications to medical diagnosis and bankruptcy prediction. He is currently a Reviewer for the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, PART B.



XUEHUA ZHAO received the Ph.D. degree from the College of Computer Science and Technology, Jilin University, in 2014. He is currently a Lecturer with the School of Digital Media, Shenzhen Institute of Information Technology. His main research interests include machine learning and data mining.

SHU-YUN ZHANG received the Ph.D. degree from the Department of Occupational and Environmental Health, Nagoya University, Japan. She is currently an Associate Professor with the School of Public Health and Management, Wenzhou Medical University, China. Her current research interests include environmental toxicology.