Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Development of novel ensemble model using stacking learning and evolutionary computation techniques for automated hepatocellular carcinoma detection

Wojciech Ksiazek [a], *Mohamed Hammad* [b], *Pawel Plawiak* [a,c,*], *U. Rajendra Acharya* [d,e,f], *Ryszard Tadeusiewicz* [g]

[a] *Department of Information and Communications Technology, Faculty of Computer Science and Telecommunications, Cracow University of Technology, ul. Warszawska 24, F-3, 31-155 Krakow, Poland*

[b] *Information Technology Department, Faculty of Computers and Information, Menoufia University, Shibin el-Kom, Menoufia, Egypt*

[c] *Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Baltycka 5, 44-100 Gliwice, Poland*

[d] *Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore*

[e] *Department of Biomedical Engineering, School of Science and Technology, Singapore School of Social Sciences, Singapore*

[f] *Department of Biomedical Informatics and Medical Engineering, Asia University, Taiwan*

[g] *Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland*

ARTICLE INFO

Article history:
Received 24 June 2020
Received in revised form
21 August 2020
Accepted 21 August 2020
Available online 24 September 2020

Keywords:
HCC
Stacking learning
Ensemble method
Machine learning
Genetic algorithm

ABSTRACT

The most common type of liver cancer is hepatocellular carcinoma (HCC), which begins in hepatocytes. The HCC, like most types of cancer, does not show symptoms in the early stages and hence it is difficult to detect at this stage. The symptoms begin to appear in the advanced stages of the disease due to the unlimited growth of cancer cells. So, early detection can help to get timely treatment and reduce the mortality rate. In this paper, we proposes a novel machine learning model using *seven* classifiers such as K-nearest neighbor (KNN), random forest, Naïve Bayes, and other *four* classifiers combined to form stacking learning (ensemble) method with genetic optimization helping to select the features for each classifier to obtain highest HCC detection accuracy. In addition to preparing the data and make it suitable for further processing, we performed the normalization techniques. We have used KNN algorithm to fill in the missing values. We trained and evaluated our developed algorithm using 165 HCC patients collected from Coimbra's Hospital and University Centre (CHUC) using stratified cross-validation techniques. There are total of 49 clinically significant features in this dataset, which are divided into *two* groups such as quantitative and qualitative groups. Our proposed algorithm has achieved the highest accuracy and F1-score of 0.9030 and 0.8857, respectively. The developed model is ready

* *Corresponding author*: Department of Information and Communications Technology, Faculty of Computer Science and Telecommunications, Cracow University of Technology, ul. Warszawska 24, F-3, 31-155 Krakow, Poland.
E-mail addresses: plawiak@pk.edu.pl, plawiak@iitis.pl, plawiak.pawel@gmail.com (P. Plawiak).

https://doi.org/10.1016/j.bbe.2020.08.007

to be tested with huge database and can be employed in cancer screening laboratories to aid the clinicians to make an accurate diagnosis.

1. Introduction

Hepatocellular carcinoma (HCC) is the most common type of liver cancer. It happens in people with chronic liver diseases, such as hepatitis C fibrosis caused by hepatitis B or C [1]. In addition, few types of liver cancers are not detected early due to the absence of symptoms in affected patients [2]. The early detection of liver cancer may reduce the treatment costs and help to save life [3].

Recently, artificial intelligence (AI) methods have helped to reduce the possible errors in the medical field and assisted to make accurate early detection of breast cancer [4,5], virus diseases [6,7], Alzheimer's disease [8,9] and cardiovascular diseases [10–17]. These AI-based techniques help to alleviate the workload of clinicians and make an accurate fast diagnosis. Hence, many researchers have proposed computer-aided diagnosis (CAD) system to detect HCC accurately [18–28].

However, the limitations of these CAD systems are as follows:

- Used smaller datasets and obtained lower performances.
- Employed fewer number of features.
- Unable to deal with the inter-patient variability.
- Requires higher execution time.

In order to overcome the above-mentioned limitations, we have introduced an ensemble learning method. Nowadays, several studies have used ensemble learning techniques in medical field to overcome the problems of conventional machine learning methods [29–33]. However, to the best of our knowledge, this is the first study to propose an ensemble classifier combining with genetic optimization for the detection of HCC.

The main contributions of this work are given below:

- Design a new machine learning model using conventional classifiers in combination as stacking learning (ensemble method) with genetic optimization of parameters and select the features in each classifier.
- K-nearest neighbor (KNN) algorithm is used to fill in the missing data to obtain higher data quality.
- This model is more accurate in detecting HCC compared to the previous models presented in the literature [18–24,45,46]. The results obtained shows that the proposed model is able to detect HCC with highest accuracy.

The remainder of this paper is organized as follows: Section 2 outlines some of the related work. Section 3 describes the proposed method with the data used in our method. Section 4 explains our experiments on the selected database. The results are shown and discussed in Section 5. Conclusion and future works are delineated briefly in Section 6.

2. Literature review

Many researchers have proposed various methods to detect and diagnose the HCC using machine learning methods. We have briefly explained them below.

- Ksiazek et al. [18], employed support vector machine (SVM) combined with *two*-level genetic optimizer to predict the HCC disease. The authors worked on 165 records from CHUC database. Their method obtained an accuracy of 88.49% to detect the HCC.
- Nayak et al. [19], performed a multi-phase analysis of computed tomography (CT) images to extract *twenty-four* features and used SVM for HCC detection. The authors worked on 40 CT images from contrast-enhanced CT dataset and reported an accuracy of 80% for detecting HCC.
- Brehar et al. [20], presented a method for HCC recognition by employing textural features coupled with adaptive boosted classifiers and reported an accuracy of 72% for HCC recognition.
- Santos et al. [21], worked on the CHUC database consisting of 165 patients. They used neural networks (NN) and logistic regression (LR) classifiers for classification and obtained an accuracy of 75.2% and 73% for NN and LR, respectively.
- Sawhney et al. [22], developed a feature selection method for cancer diagnosis using firefly algorithm (FFA) with random forest classifier. They reported an accuracy of 83.5% using same CHUC database.
- Aonpong et al. [24], employed least absolute shrinkage and selection operator (LASSO) regression for features selection for early recurrence prediction of HCC. They obtained 89.18% classification accuracy using SVM and decision tree classifiers.
- Hammad et al. [45], presented two methods to classify HCC data. The two methods are: supervised missing feature completion method and weight-based feature reduction method. They used 23 conventional and ensemble classifiers for classification such as Logistic Regression, Discriminant (Linear and Quadratic), SVM (Linear, Quadratic, Cubic, Fine Gaussian, Medium Gaussian and Coarse Gaussian), KNN (Fine, Medium, Coarse, Cosine, Cubic and Weighted) and other 8 classifiers. They obtained the best accuracies using NCA and relief of 92.12% and 83.03%, respectively.
- Zheng et al. [46], developed a machine learning based method to diagnose HCC. They removed relative expression orderings (REOs) using minimum redundancy maximum relevance (mRMR). They obtained sensitivity, specificity, and an AUC of 91.93%, 100%, and 95.97%, respectively.

- Acharya et al. [58], presented a hybrid system using *three* algorithms, such as linear discriminant analysis (LDA), SVM and GA. They used LDA to reduce the number of features and SVM for classification. They used GA to optimize the model and obtained an accuracy of 90.30%, sensitivity of 82.25% and specificity of 96.07%.

Table 10 summarizes the previous methods used for automated detection of HCC using machine learning methods.

The previous methods have many limitations and we try to overcome most of these limitations by proposing a novel and robust method that combine the ensemble learning with genetic algorithm (GA) for the detection of HCC. This two-layered model built using different classifiers with the Nu-SVC as meta classifier. Additionally, the use of the algorithm of the k nearest neighbors made it possible to precise fill in the missing values. The whole has been optimized using evolutionary computations. This combination helped to obtain high accuracy in detecting the HCC.

3. Materials and method

The material and the methodology used in our proposed work are discussed in this section. In addition, we discussed the HCC database used to implement our method.

3.1. *HCC dataset*

In this paper, we implemented our method using 165 HCC patients collected from CHUC [21]. There are total of 49 clinical features in this database, which are being used by the clinicians to make an accurate diagnosis. These features are divided into two groups: the first group with 23 features are called the quantitative group and the second group with 26 features are called the qualitative group. This database is described in detail in Table 1 10.2% of the whole data are missing in the database. This table shows the type and scale of each feature. The statistics of the feature is presented by the *mean* and range of each feature. Eight patients from 165 total patients have whole information in all fields. The dataset divided into two classes-***dead*** class (label as 0) with 63 patients and ***alive*** class (label as 1) with 102 patients.

The Supplementary materials present histograms for the following characteristics: age, gender, class.

3.2. *Methodology*

This section describes the different stages of the proposed methodology as shown in Fig. 1 with missing value, preprocessing, GA (GA features selection, GA parameter optimization) and classification steps. The main parts of each stage are briefly explained below.

3.2.1. *Missing values*

In this step, the Python imputations library Impyute [34] is used. It is designed to estimate missing values in the HCC database.

We have used the K-nearest neighbor (KNN) algorithm [48] to fill the missing values. KNN can keep the original input data distribution by selecting suitable K (in our case, K = 5) value. Table 2 shows the configuration of KNN parameters.

The nearest neighbor's algorithm makes it possible to estimate missing data based on several closest samples. This enables a more accurate estimate than conventional methods to fill in missing values such as mode, median or average, which are examples of global estimation [23]. Obtaining better quality data leads to the preparation of an effective classification model.

3.2.2. *Preprocessing and normalization*

After completing the missing values, the normalization method is applied. During normalization, the data is reorganized so that we can utilize the data for further analysis. The main goal of normalization is to group the data together and get rid of any duplicate data that might appear within the database. The common normalization methods are min-max, Z-score, etc. [35]. The optimal normalization method depends on the data to be normalized. We found that used min-max method is most suitable for HCC data. Therefore, we used min-max scaler normalization and data is rescaled to the range [0,1] in this step.

3.2.3. *GA for feature selection and parameters optimization*

GA is a method of optimization and research, which can be classified as one of the methods of evolutionary algorithms [36,55]. It is used to find exact or approximate solutions that optimize, genetic algorithms classified as global research heuristics, and it is known as an evolutionary computation that uses technology inspired by evolutionary biology. Such as heritability, mutation, selection, and crossover. In this paper, *two*-layer GA is proposed for feature selection and parameter optimization of all classifiers. In the first layer, GA [36] is used with *seven* classic classifiers namely KNN [48], random forest [49], Naïve Bayes [50], SVC [51], NuSVM [52], logistic regression [53] and linear discriminant analysis (LDA) [54] are used to optimize their parameters and select the features. In the second layer, GA is used with meta-classifier (NuSVM [52]) to optimize only its parameters and perform the classification.

For feature selection using GA, the most widely used binary string encoding [37] yielded 1 when the feature is selected and 0 if it is not selected. Finally, we optimized the parameters of each classifier using GA. Table 3 shows the details of GA parameters used in our study. Table 4 shows the optimization parameters in each classifier used using GA.

3.2.4. *Ensemble learning (classification) and stacking learning*

Ensemble learning uses more classifiers to obtain better predictive performance than a single classifier. Stacking learning is an ensemble method that combines multiple classifiers via a meta-classifier [38]. In this work, we employed stacking learning as a classifier to classify the meta-features to obtain the final class.

Classifiers from the first layer (genetic optimization of parameters + genetic selection of features) return the probability of belonging to a class (meta-features). In the second layer, these meta-features are the input of the meta-classifier (Nu -SVM with genetic optimization of parameters). Finally, the output of the classifier can be 1 (in case of a patient survives) or 0 (in case of a patient died).

Table 1 – Description of the clinical features of the HCC database.

Clinical features	Features type	Features scale	Mean	Range	Missing values (%)
Gender	Qualitative	Dichotomous	1	0,1	0
Symptoms			1		10.9
Alcohol			1		0
HBsAg			0		10.3
HBeAg			0		23.6
HBcAb			0		14.6
HCVAb			0		5.5
Cirrhosis			1		0
Endemic countries			0		23.6
Smoking			1		24.9
Diabetes			0		1.8
Obesity			0		6.1
Hemochromatosis			0		13.9
AHT			0		1.8
CRI			0		1.2
HIV			0		8.5
NASH			0		13.3
Esophageal varices			1		31.5
Splenomegaly			1		9.1
Portal hypertension			1		6.7
Portal vein thrombosis			0		1.8
Liver metastasis			0		2.4
Radiological hallmark			1		1.2
Performance status		Ordinal	0	0, 1, 2, 3, 4	0
Encefalopathy			1	1, 2, 3	0.6
Ascites			1	1, 2, 3	1.2
Age at diagnosis	Quantitative	Ratio	64.69	20–93	0
Grams/day			71.01	0–500	29.1
Packs/year			20.46	0–510	32.1
INR			1.42	0.84–4.82	2.4
AFP			19,299.95	1.2–1,810,346	4.9
Hemoglobin			12.88	5–18.7	1.8
MCV			95.12	69.5–119.6	1.8
Leukocytes			1473.96	2.2–13,000	1.8
Platelets			113,206.44	1.71–459,000	1.8
Albumin			3.45	1.9–4.9	3.6
Total Bil			3.09	0.3–40.5	3
ALT			67.09	11–420	2.4
AST			69.38	17–553	1.8
GGT			268.03	23–1575	1.8
ALP			212.21	1.28–980	1.8
TP			8.96	3.9–102	6.7
Creatinine			1.13	0.2–7.6	4.2
Number of nodules			2.74	0–5	1.2
Major dimension			6.85	1.5–22	12.1
Dir. bil			1.93	0.1–29.3	26.7
Iron (mcg/dL)			85.599	0–244	47.88
Sat			37.029	0–126	48.48
Ferritin			438.998	0–2230	48.48

Our stacking classifier consists of 7 classifiers in the first layer and a meta classifier in the second layer. The classifiers from the first layer return the probabilities that go into the meta classifier input. The first layer uses genetic optimization of parameters and features. In the case of the meta-classifier, only genetic optimization of parameters was used.

3.2.5. *Cross validation*

Stratified K-Folds cross-validation (k = 5) approach [39] is used in this work to ensure that relative class frequencies are approximately preserved in each training and validation fold. In this approach, the testing and training sets are created by randomly selecting HCC data separately for each class, while maintaining the proportions between classes. Cross-validation is carried out on the entire available database.

3.2.6. *Evaluation metrics*

To evaluate the performance of our method, we used *two* basic metrics: accuracy and F1-score [44].

Our metrics are computed based on the confusion matrix. Table 5 presents a typical example of the confusion matrix.

Accuracy: It is given by Eq. (1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

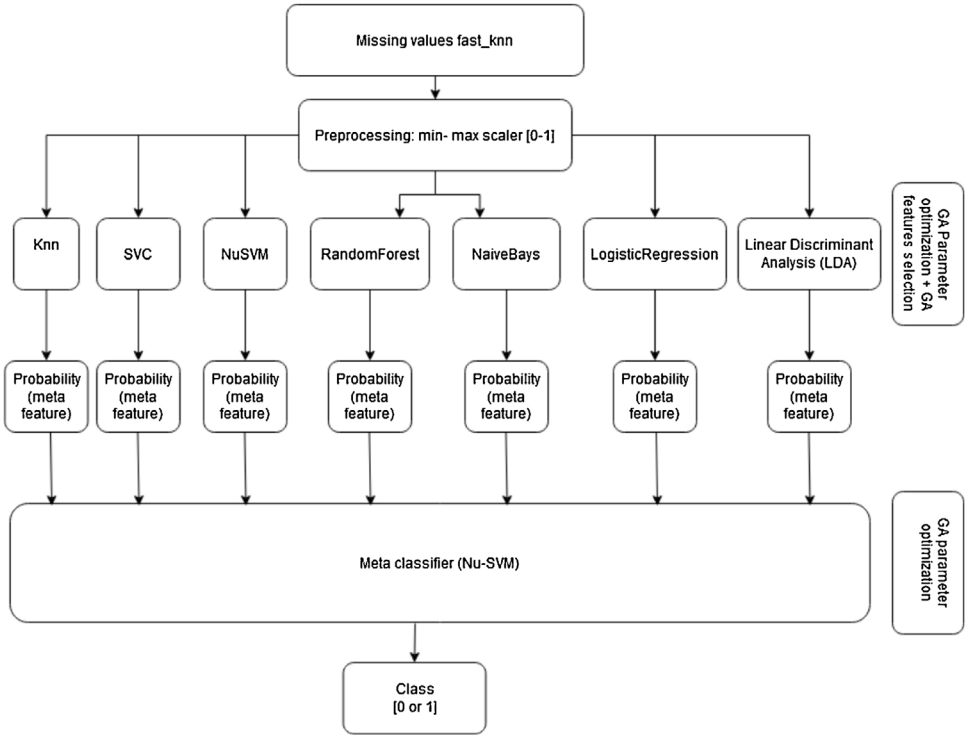

Fig. 1 – Proposed model used for the automated detection of HCC.

F1-score: It is obtained from recall and precision metrics. It is given by Eq. (4).

$$Recall = \frac{TP}{TP + FN} \quad (2) \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \quad (3) \tag{3}$$

$$F1\_Score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{4}$$

4. Results

The machine learning model is developed using Python 3.7. The following libraries are used:

a) Pandas [40] — loading data
b) Impyute [34] — estimating missing values
c) Sklearn [41] — use of based classifiers
d) StackingClassifier [42] — combining classifiers into ensemble (stacking learning)
e) Deap [43] — parameter optimization and selection of features using genetic algorithms.

Specification of the computer on which this work is performed are as follows:

a) Processor: Intel Core i5-7300HQ 2.5 GHz
b) RAM: 16 GB

Table 2 – Configuration parameters of KNN used in this study.

Parameter	Value
k (number of neighbors)	5
p (metrics)	2 (Euclidean)

Table 3 – Typical GA parameters used in our work.

Parameter	Value
Selection algorithm	Tournament selection – size of tournament: 3
Crossover method	Two-point crossover
Mutation method	One-point mutation (own algorithm)
Probability of crossover	0.8
Probability of mutation	0.8
Size of population	100
Number of epoch	200
Elitist strategy	1%
Fitness function	Accuracy or F1-score

Table 4 – Optimized parameters used in each classifier with GA.

Classifier	Parameters	Value
KNN	K	[1–10]
	Metric	["euclidean","manhattan","chebyshev"]
	Weight	["uniform","distance"]
Random Forest	Features selection only	–
Naïve Bayes	Features selection only	–
SVC	Kernel	["linear","rbf", "poly","sigmoid"]
	C	[0.1–100]
	Degree	[1–5]
	Gamma	[0.001–5]
NuSVM	Kernel	["linear","rbf", "poly","sigmoid"]
	Nu	[0.001–0.5]
	Degree	[1–5]
	Gamma	[0.001–5]
Logistic Regression	C	[1–100]
	Max iter	[1–2000]
LDA	Features selection only	–

Table 5 – Typical confusion matrix.

Actual	Predicted	
	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Table 6 – Confusion matrix for first model.

Actual	Predicted	
	Survive	Dead
Survive	97	5
Dead	11	52

4.1. *First experiment — model with accuracy as fitness function*

In the first step of the experiment the model described in Section 3.2.4. was used. The selection of features for each classifier and optimization of its parameters has been made. Accuracy was selected as a fitness function of genetic algorithm. A formula is provided in Section 3.2.6. The model was developed using 165 HCC patients. The database has been described in detail in Section 3.1.

The best model achieved 90.30% accuracy rate on testing set. Table 6 shows confusion matrix for this model.

It can be noted from Table 6 that, 11 dead records were detected as survived. However, 5 survive records were classified as dead samples.

Fig. 2 shows the variation of F1-score for various epochs of the genetic algorithm.

Fig. 2 shows that with successive epochs of the genetic algorithm the accuracy of the classifier increases. Over 200 epochs, the model performance has increased by about 15%.

Table 7 shows the parameters of the classifiers and selected features for the best model.

Table 7 contains the details of all classifiers included in the model. There were 7 classifiers returning the probability of class (KNN, SVC, NuSVM, Logistic Regression, Random Forest, LDA, Naïve Bayes) and the meta-classifier NuSVM. The Table contains information about the parameters of individual classifiers. Moreover, information about the selected features has been

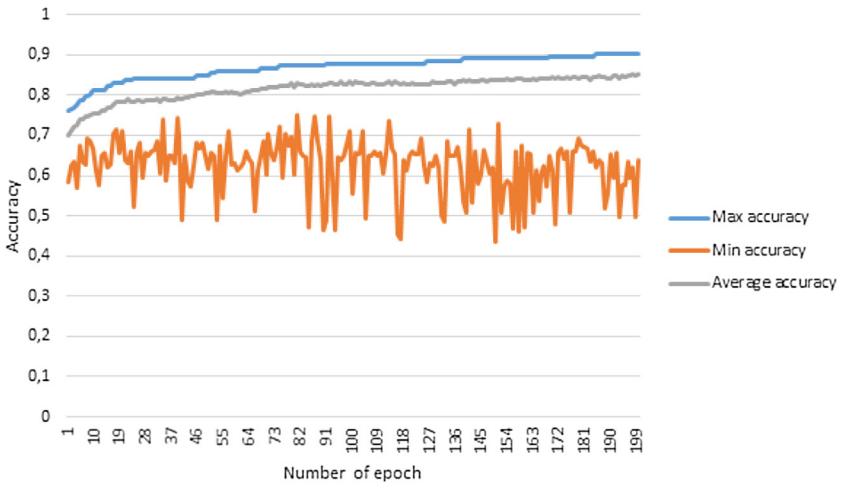

Fig. 2 – Variation of accuracy for various epochs of GA.

Table 7 – Analysis of the best individual for accuracy.

Classifier	Optimized values
KNN	3, 'chebyshev', 'uniform' (k, metric, weight)
SVC	'sigmoid', 26.42686497495269, 2.3015795072861573, 0.1106544214142905 (kernel, C, degree, gamma)
NuSVM	'linear', 0.4969215671722941, 4, 0.5410943558906997 (kernel, Nu, degree, gamma)
Logistic Regression	88.9514224188211, 129 (C, max_iter)
Meta classifier NuSVM	'rbf', 0.313013089448576, 2, 3.006554536476033 (kernel, Nu, degree, gamma)
Features for KNN	[0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0] Selected features: Symptoms, Alcohol, HCVAb, Diabetes, Hemochromatosis, CRI, NASH, Esophageal varices, Splenomegaly, Portal vein thrombosis, Performance status, Encefalopathy, Packs/year, AFP, Hemoglobin, MCV, Leukocytes, Albumin, ALT, GGT, Number of nodules, Dir. bil, Iron
Features for SVC	[1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0] Selected features: Gender, HBsAg, HBeAg, HCVAb, Hemochromatosis, AHT, CRI, HIV, Portal vein thrombosis, Liver metastasis, Radiological hallmark, Grams/day, Packs/year, AFP, Platelets, Albumin, ALT, Creatinine, Number of nodules, Sat
Features for NuSVM	[1,0,0,0,1,1,0,0,1,1,1,0,1,1,0,1,1,1,1,1,1,1,1,0,0,1,0,0,0,0,1,0,0,0,0,1,1,1, 1,1,1,0,1,1,1,0,1,1,1] Selected features: Gender, HBeAg, HBcAb, Endemic countries, Smoking, Diabetes, Hemochromatosis, AHT, HIV, NASH, Esophageal varices, Splenomegaly, Portal hypertension, Portal vein thrombosis, Liver metastasis, Radiological hallmark, Ascites, AFP, Albumin, Total Bil, ALT, AST, GGT, ALP, Creatinine, Number of nodules, Major dimension, Iron (mcg/dL), Sat, Ferritin
Features for Random Forest	[1,0,1,1,0,1,0,0,1,1,0,0,1,1,0,0,0,0,0,0,1,1,1,1,0,1,0,1,0,0,1,1,1,1,0,0,1,1,1,1, 1,0,0,0,0,0,0,0,1] Selected features: Gender, Alcohol, HBsAg, HBcAb, Endemic countries, Smoking, Hemochromatosis, AHT, Portal vein thrombosis, Liver metastasis, Radiological hallmark, Performance status, Ascites,Grams/day, AFP, Hemoglobin, MCV, Leukocytes, Total Bil, ALT, AST, GGT, ALP, Ferritin
Features for Logistic Regression	[1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0] Selected features: Gender, Symptoms, Alcohol, HBeAg, HBcAb, Endemic countries, Smoking, Obesity, AHT, CRI, Splenomegaly, Portal vein thrombosis, Performance status, INR, AFP, Hemoglobin, Leukocytes, Platelets, AST, ALP, TP, Number of nodules, Sat
Features for LDA	[1,0,0,0,1,1,1,1,1,1,1,1,0,1,1,0,0,1,1,1,0,1,1,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0, 1,1,0,0,1,1,1] Selected features: Gender, HBeAg, HBcAb, HCVAb, Cirrhosis, Endemic countries, Smoking, Diabetes, Obesity, Hemochromatosis, CRI, HIV, Splenomegaly, Portal hypertension, Portal vein thrombosis, Radiological hallmark, Performance status, Grams/day, Hemoglobin, MCV, Leukocytes, ALT, AST, Creatinine, Number of nodules, Iron (mcg/dL), Sat, Ferritin
Features for Naïve Bayes	[0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0] Selected features: Symptoms, Alcohol, HCVAb, Diabetes, Hemochromatosis, NASH, Esophageal varices, Splenomegaly, Portal vein thrombosis, Performance status, Encefalopathy, Packs/year, AFP, Hemoglobin, MCV Leukocytes, Albumin, ALT, GGT, Number of nodules, Dir. bil, Iron (mcg/dL)

added also in this Table, where 1 refers to feature selected and 0 refers to feature rejected. The genetic algorithm chose 23 features for KNN, 21 for SVC, 30 for NuSVM, 24 for Random Forest, 23 for Logistic Regresion, 29 for LDA, 23 for Naïve Bayes.

The random forest is tested with the following parameters:

a) criterion: entropy
b) n_estimators: 100

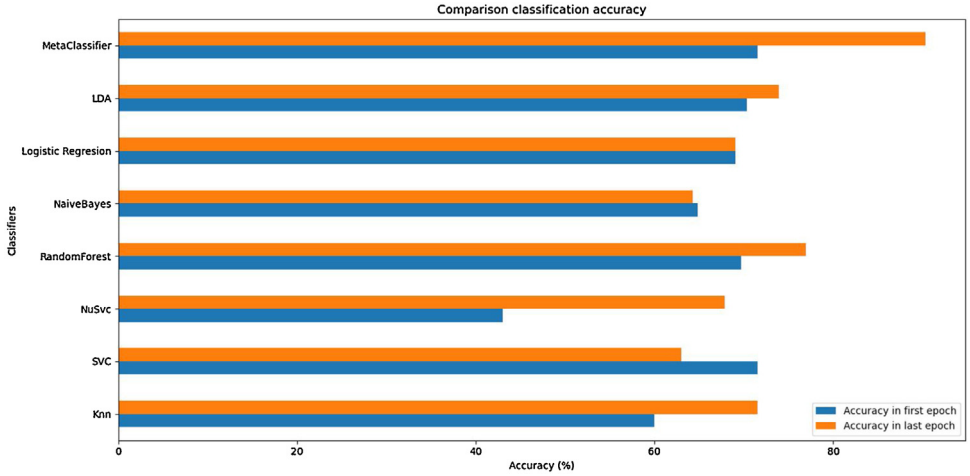

Fig. 3 – Comparison of various classifiers based on accuracy (%).

LDA and Naïve Bayes used the default parameters (Naïve Bayes: var_smoothing = 1e-09, LDA: solver = 'svd' and tol = 0.0001).

Fig. 3. shows accuracy of each classifier.

It can be noted from Fig. 3 that, the accuracy of individual classifiers is not very high. In few cases, individual classifiers obtained lower accuracies in the last epochs than in the beginning. This may be because the genetic algorithm aims to ensure that the meta classifier obtains the highest accuracy and not individual classifiers.

Only the meta-classifier, learning from the probability returned by a single classifier, achieved a high accuracy of 90.30%.

4.2. *Second experiment – model with F1-score as fitness function*

In the second experiment, the model proposed in Section 3.2.4 was again used. However, modifications were made to the fitness function optimized by the genetic algorithm. This time the F1-score described in Section 3.2.6 as a formula 4 was used. Parameters were again optimized, and features selected for all classifiers were included in the model.

The best F1-score of 88.57% is achieved on testing set. Table 8 shows the confusion matrix for these calculations.

According to Table 8, it can be seen that7 dead records were detected as survived. However, 11 survive records were classified as dead samples.

Fig. 4 shows the variation of F1-score obtained for various epochs of the genetic algorithm.

Table 8 – Confusion matrix obtained using F1-score.

Actual	Predicted	
	Survive	Dead
Survive	91	11
Dead	7	56

It can be noted from Fig. 4 that. it can see the coincidence of the genetic algorithm over time. The evolution process improved the classifier score by 15% compared to the results of the first epoch.

Table 9 shows the parameters of the classifiers and selected features for model with F1-score as fitness function.

In Table 9, we can find the parameters of all classifiers, as well as the features selected by individual classifiers (1 - selected feature, 0 - rejected feature). Genetic algorithms have enabled each classifier to select features to achieve the highest results. The genetic algorithm chose 21 features for KNN, 23 for SVC, 26 for NuSVM, 21 for Random Forest, 21 for Logistic Regression, 24 for LDA, 21 for Naïve Bayes. Finally, the meta-classifier makes the final classification based on the probabilities returned by individual classifiers.

The random forest was tested with the following parameters:

c) criterion: entropy
d) n_estimators: 100

LDA and Naïve Bayes used the default parameters (Naïve Bayes: var_smoothing = 1e-09, LDA: solver = 'svd' and tol = 0.0001).

Fig. 5 shows the comparison of classification performance of various classifiers obtained based on F1-score. In this experiment also we can see that the genetic algorithm maximizes the score obtained by the meta-classifier. The results obtained by single models are not increased gradually in the next epochs.

5. Discussion

Table 10 shows the comparison of state-of-art techniques developed for the automated detection of HCC disease using the same database (CHUC).

An important observation is that, we have achieved higher accuracy and F1-score compared to the reported previous works using the same database. To the best of our knowledge,

Fig. 4 – Variation of F1-score obtained for various number of epochs of GA.

Table 9 – Summary of parameters and selected features used for various classifiers.

Classifier	Optimized values
KNN	3, 'chebyshev', 'uniform' (k, metric, weight)
SVC	'linear', 50.55387072804635, 1.7947967836849261, 4.735376436182018 (kernel, C, degree, gamma)
NuSVM	'linear', 0.2640102766647352, 3, 1.530273475303261 (kernel, Nu, degree, gamma)
Logistic Regression	65.03920362515971, 53 (C, max_iter)
Meta classifier NuSVM	'poly', 0.46878689623925784, 5, 4.819446218603088 (kernel, Nu, degree, gamma)
Features for KNN	[0,1,0,1,1,1,1,0,1,0,1,0,0,1,0,0,1,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,1,1,1,0,0,0,1,0, 0,1,0,1,1,0,0,0,1] Selected features: Symptoms, HBsAg, HBeAg, HBcAb, HCVAb, Endemic countries, Diabetes, AHT, NASH, Ascites, Age at diagnosis, Grams/day, Packs/year, MCV, Leukocytes, Platelets, AST, TP, Number of nodules, Major dimension, Ferritin
Features for SVC	[1, 0, 1, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0] Selected features: Gender, Alcohol, HBsAg, HBeAg, HBcAb, Cirrhosis, Diabetes, Obesity, Hemochromatosis, AHT, HIV, Esophageal varices, Portal vein thrombosis, Radiological hallmark, Performance status, Encefalopathy, Age at diagnosis, Packs/year, Hemoglobin, Leukocytes, Total Bil, ALT, Creatinine
Features for NuSVM	[0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0] Selected features: Symptoms, HBsAg, HCVAb, Smoking, Diabetes, AHT, CRI, HIV, Portal vein thrombosis, Performance status, Ascites, Age at diagnosis, Grams/day, Packs/year, INR, Hemoglobin, MCV, ALT, ALP, TP, Number of nodules, Major dimension, Dir. Bil, Iron (mcg/dL), Sat
Features for Random Forest	[1,0,0,0,0,0,1,1,0,0,0,1,1,1,0,0,0,0,1,1,1,1,0,1,0,1,0,1,0,1,1,0,1,0,1,0,0,0,0,0, 1,0,1,0,1,0,0,0,0,1,1] Selected features: Gender, HCVAb, Cirrhosis, Obesity, Hemochromatosis, AHT, Splenomegaly, Portal hypertension, Portal vein thrombosis, Liver metastasis, Performance, Ascites, Grams/day, INR, AFP, MCV, AST, ALP, Creatinine, Sat, Ferritin
Features for Logistic Regression	[0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0] Selected features: Symptoms, Alcohol, HBcAb, Diabetes, AHT, NASH, Portal hypertension, Portal vein thrombosis, Age at diagnosis, INR, AFP, Hemoglobin, MCV, Leukocytes, Albumin, Total Bil, ALP, Major dimension, Dir. Bil, Iron (mcg/dL), Sat
Features for LDA	[1,1,0,1,1,1,1,0,1,1,0,0,0,1,0,0,0,0,1,0,1,0,0,0,1,0,0,0,0,1,0,1,0,1,0,0,1,1,1, 0,1,1,1,0,1,0,1,0,1]

Table 9 (*Continued*)

Classifier	Optimized values
Features for Naïve Bayes	Selected features: Gender, Symptoms, HBsAg, HBeAg, HBcAb, HCVAb, Endemic countries, Smoking, AHT, Splenomegaly, Portal vein thrombosis, Encefalopathy, INR, Hemoglobin, Leukocytes, Total Bil, ALT, AST, ALP, TP, Creatinine, Major dimension, Iron (mcg/dL), Ferritin [0,1,0,1,1,1,1,0,1,0,1,0,0,1,0,0,1,0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,0,1,1,1,0,0,0,1,0,0,1,0, 1,1,0,0,0,1] Selected features: Symptoms, HBsAg, HBeAg, HBcAb, HCVAb, Endemic countries, Diabetes, AHT, NASH, Ascites, Age at diagnosis, Grams/day, Packs/year, MCV, Leukocytes, Platelets, AST, TP, Number of nodules, Major dimension, Dir. Bil, Ferritin

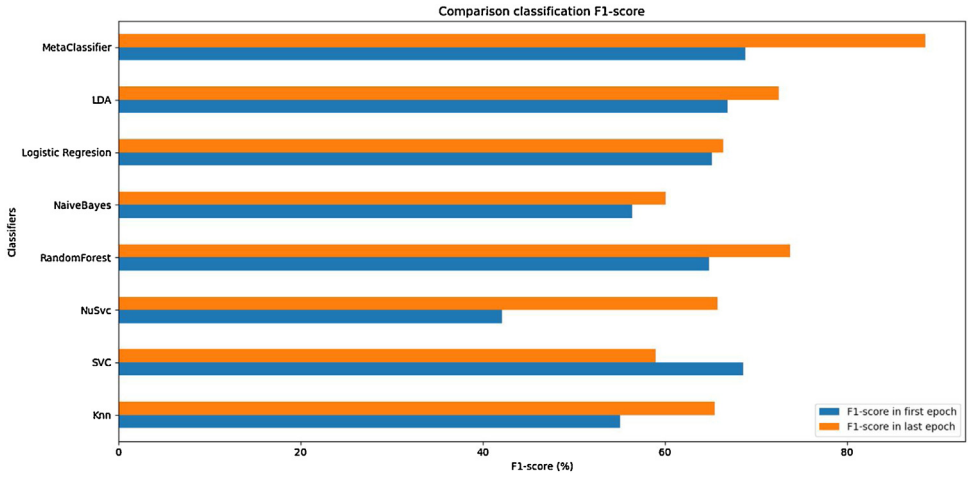

Fig. 5 – Comparison of classification performance of various classifiers obtained based on F1-score.

Table 10 – Comparison of state-of-art techniques developed for the automated detection of HCC disease using the same database (CHUC).

Study	Algorithm	Accuracy	F1-score
Ksiazek et al. [18]	SVC + GA	0.8849	0.8762
Nayak et al. [19]	SVM with RBF kernel	0.869	0.399
Brehar et al. [20]	SVM + AdaBoost	0.720	N.A.
Santos et al. [21]	NN + Augmented sets approach	0.7519 ± 0.0105	0.6650 ± 0.0182
Sawhney et al. [22]	BFA + RF	0.835	N. A.
Aonpong et al. [24]	LASSO + SVM-RFE	0.8918	N.A.
Tuncer and Ertam [46]	reliefF + LDA	0.8303	0.8202
	NCA + FGSVM	0. 9212	0.9161
Ali et al. [59]	LDA–GA–SVM (with linear and RBF kernel)	0.9030	N.A.
This study	StackingGA	0.9030	0.8857

this is the first study to employ the ensemble learning technique based on the combination of stacking learning with genetic optimization for automatic detection of HCC.

The advantages of the proposed method are as follows:

a) Achieved highest accuracy and F1-score.
b) Used various types of classifiers to obtain stable results.
c) Employed evolutionary algorithm to select the best parameters and features to develop the stable optimum performing model.
d) Used KNN algorithm to fill the missing data to obtain high quality data.

The disadvantages of the proposed method are given below:

a) Requires long model training time.
b) Developed model is more complex.
c) Need to more data to develop a more stable and accurate model. In the future, we plan to validate our system with more databases.

As shown in Table 10, our proposed model achieved the best results as compared to the state-of-art techniques. Our innovative combination of genetic algorithms with ensemble

learning – stacking has yielded the best performance in detecting the HCC disease. This is the first work proposed using ensemble learning to detect HCC. In [18–20,45], the authors obtained good performance using more number of features. However, the proposed method achieved the highest performance using fewer features. In the algorithm proposed in [21], has obtained lower performance and hence may not be suitable for real world applications. Other works in Refs. [18–22,24,45] have obtained need to be tested with more data. The method in Ref. [58] obtained the same accuracy as our proposed method, however, our proposed method is more robust and less complex.

The proposed model is presented in the Fig. 1 which consists of 7 classifiers and a meta-classifier. The classifiers described in the ensemble model is presented in detail in Table 4. Genetic optimization of classifier parameters included in the first layer and selection of their features, as well as optimization of the meta-classifier parameters provided excellent results.

In Figs. 3 and 5 we can see the result of the individual classifiers and the meta classifier. It can be seen that the result of individual classifiers is much lower than the meta-classifier (90.30% for accuracy and 88.57 for F1-score). We can see that the individual classifier focuses not on its own results, but on the end result of the meta classifier. Detailed results for the meta classifiers are presented in the confusion matrices (Tables 6 and 8).

Additionally, it is worth emphasizing the importance of genetic algorithms. In Figs. 2 and 4, we can see the convergence of methods and significant (several percentages) improvement in the results. Appropriately selected parameters of the genetic algorithm (including selection algorithm, crossover, number of epochs, population size) have yielded significantly improved performance as presented in Table 3.

Genetic algorithms help to select clinically significant features for individual classifiers. It can be noted from Tables 7 and 9 that, data reduction has resulted in the reduction of number of features within individual classifiers is about 50%. As a result, the model is less complex than the classical classifiers using all the features.

Our ensemble model obtained highest classification performance, but the model is complex and takes long learning time. This ensemble can still be developed by adding more new classifiers or by modifying the existing ones. Perhaps joining several neural networks may also improve the classification performance. In addition, each classifier of ensemble can perform different data preprocessing. This approach can provide greater variety of classifiers, which is the key aspect of ensemble learning.

We intend to use deep learning techniques in the future for the detection of HCC to increase the diagnosis performance using big data. We also plan to use other ensemble learning methods namely voting classifier, deep stacking classifier or boosting algorithm using our dataset [47]. Finally, we plan to try the proposed method for detecting other types of diseases such as liver diseases [56,57].

6. Conclusions

In this study, a novel algorithm based on the combination of stacking learning with genetic optimization for automated detection of HCC disease is proposed. We have used different classifiers with and without feature selection on the HCC dataset. This model enabled more accurate detection of HCC than previous models presented in the literature. Our proposed method obtained an overall accuracy of 0.9030 and F1-score of 0.8857. Therefore, our system is an effective tool to perform an accurate and consistent diagnosis of HCC.

CRediT authorship contribution statement

Wojciech Ksiazek: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Mohamed Hammad:** Writing - original draft, Writing - review & editing. **Pawel Plawiak:** Conceptualization, Methodology, Formal analysis, Writing - review & editing, Supervision. **U. Rajendra Acharya:** Writing - review & editing. **Ryszard Tadeusiewicz:** Writing - review & editing.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.bbe.2020.08.007.

REFERENCES

[1] Fung J, Lai CL, Yuen MF. Hepatitis B and C virus-related carcinogenesis. Clin Microbiol Infect 2009;15(11):964–70.
[2] Bruix J, Reig M, Sherman M. Evidence-based diagnosis, staging, and treatment of patients with hepatocellular carcinoma. Gastroenterology 2016;150(4):835–53.
[3] Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, et al. The case for early detection. Nat Rev Cancer 2003;3(4):243–52.
[4] Raghavendra U, Gudigar A, Rao TN, Ciaccio EJ, Ng EYK, Acharya UR. Computer-aided diagnosis for the identification of breast cancer using thermogram images: a comprehensive review. Infrared Phys Technol 2019;102103041.
[5] Khan SU, Islam N, Jan Z, Din IU, Khan A, Faheem Y. An e-Health care services framework for the detection and classification of breast cancer in breast cytology images as an IoMT application. Future Gener Comput Syst 2019;98:286–96.
[6] Santosh KC. AI-driven tools for coronavirus outbreak: need of active learning and cross-population train/test models on multitudinal/multimodal data. J Med Syst 2020;44(5):1–5.
[7] Yao Z, Li J, Guan Z, Ye Y, Chen Y. Liver disease screening based on densely connected deep neural networks. Neural Netw 2020;123:299–304.
[8] Wei JKE, Jahmunah V, Pham TH, Oh SL, Ciaccio EJ, Acharya UR, et al. Automated detection of Alzheimer's disease using bi-directional empirical model decomposition. Pattern Recognit Lett 2020;135:106–13.
[9] Johansson P, Almqvist EG, Bjerke M, Wallin A, Johansson JO, Andreasson U, et al. Reduced cerebrospinal fluid concentration of apolipoprotein AI in patients with Alzheimer's disease. J Alzheimer Dis 2017;59(3):1017–26.

[10] Abdar M, Ksiazek W, Acharya UR, Tan RS, Makarenkov V, Plawiak P. A new machine learning technique for an accurate diagnosis of coronary artery disease. Comput Methods Programs Biomed 2019;179104992.
[11] Hammad M, Maher A, Wang K, Jiang F, Amrani M. Detection of abnormal heart conditions based on characteristics of ECG signals. Measurement 2018;125:634–44.
[12] Alghamdi A, Hammad M, Ugail H, Abdel-Raheem A, Muhammad K, Khalifa HS, et al. Detection of myocardial infarction based on novel deep transfer learning methods for urban healthcare in smart cities. Multimed Tools Appl 2020;1–22.
[13] Amrani M, Hammad M, Jiang F, Wang K, Amrani A. Very deep feature extraction and fusion for arrhythmias detection. Neural Comput Appl 2018;30(7):2047–57.
[14] Kandala RN, Dhuli R, Plawiak P, Naik GR, Moeinzadeh H, Gargiulo GD, et al. Towards real-time heartbeat classification: evaluation of nonlinear morphological features and voting method. Sensors 2019;19(23):5079.
[15] Nasarian E, Abdar M, Fahami MA, Alizadehsani R, Hussain S, Basiri ME, et al. Association between work-related features and coronary artery disease: a heterogeneous hybrid feature selection integrated with balancing approach. Pattern Recognit Lett 2020;133:33–40.
[16] Tuncer T, Dogan S, Plawiak P, Acharya UR. Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals. Knowledge Based Syst 2019;186104923.
[17] Zomorodi-moghadam M, Abdar M, Davarzani Z, Zhou X, Plawiak P, Acharya UR. Hybrid particle swarm optimization for rule discovery in the diagnosis of coronary artery disease. Expert Syst 2019e12485.
[18] Ksiazek W, Abdar M, Acharya UR, Plawiak P. A novel machine learning approach for early detection of hepatocellular carcinoma patients. Cogn Syst Res 2019;54:116–27.
[19] Nayak A, Kayal EB, Arya M, Culli J, Krishan S, Agarwal S, et al. Computer-aided diagnosis of cirrhosis and hepatocellular carcinoma using multi-phase abdomen CT. Int J Comput Assist Radiol Surg 2019;14(8):1341–52.
[20] Brehar R, Mitrea D, Nedevschi S, Lupsor MP, Rotaru M, Badea R. Hepatocellular carcinoma recognition in ultrasound images using textural descriptors and classical machine learning. 2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP); 2019. pp. 491–7. IEEE.
[21] Santos MS, Abreu PH, García-Laencina PJ, Simão A, Carvalho A. A new cluster-based oversampling method for improving survival prediction of hepatocellular carcinoma patients. J Biomed Inform 2015;58:49–59.
[22] Sawhney R, Mathur P, Shankar R. A firefly algorithm based wrapper-penalty feature selection method for cancer diagnosis. International Conference on Computational Science and its Applications; 2018. pp. 438–49. Springer, Cham.
[23] Malarvizhi MsR, Thanamani AS. K-nearest neighbor in missing data imputation. Int J Eng Res Dev 2012;5 (November (1)):05–7.
[24] Aonpong P, Chen Q, Iwamoto Y, Lin L, Hu H, Zhang Q, et al. Comparison of machine learning-based radiomics models for early recurrence prediction of hepatocellular carcinoma. J Image Graph 2019;7(4).
[25] Shen J, Qi L, Zou Z, Du J, Kong W, Zhao L, et al. Identification of a novel gene signature for the prediction of recurrence in HCC patients by machine learning of genome-wide databases. Sci Rep 2020;10(1):1–9.
[26] Liao H, Xiong T, Peng J, Xu L, Liao M, Zhang Z, et al. Classification and prognosis prediction from histopathological images of hepatocellular carcinoma by a fully automated pipeline based on machine learning. Ann Surg Oncol 2020;1–11.
[27] Das A, Acharya UR, Panda SS, Sabut S. Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques. Cognitive Systems Research 2019;54:165–75.
[28] Demir FB, Tuncer T, Kocamaz AF, Ertam F. A survival classification method for hepatocellular carcinoma patients with chaotic Darcy optimization method based feature selection. Med Hypotheses 2020;139109626.
[29] Plawiak P, Acharya UR. Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. Neural Comput Appl 2020;32:11137–61. Springer.
[30] Brunese L, Mercaldo F, Reginelli A, Santone A. An ensemble learning approach for brain cancer detection exploiting radiomic features. Comput Methods Programs Biomed 2020;185105134.
[31] Plawiak P. Novel genetic ensembles of classifiers applied to myocardium dysfunction recognition based on ECG signals. Swarm Evol Comput 2018;39:192–208.
[32] Lyu Q, Shan H, Wang G. MRI super-resolution with ensemble learning and complementary priors. IEEE Trans Comput Imaging 2020;6:615–24.
[33] Shakeel PM, Tolba A, Al-Makhadmeh Z, Jaber MM. Automatic detection of lung cancer from biomedical data set using discrete AdaBoost optimized ensemble learning generalized neural networks. Neural Comput Appl 2020;32 (3):777–90.
[34] Impyute, 2017. https://impyute.readthedocs.io/en/latest/. [Accessed 7 April 2020].
[35] García S, Luengo J, Herrera F. Data preprocessing in data mining. Cham, Switzerland: Springer International Publishing; 2015. p. 195–243.
[36] Plawiak P, Abdar M, Acharya UR. Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring. Appl Soft Comput 2019;84105740.
[37] Plawiak P. Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system. Expert Syst Appl 2018;92:334–49.
[38] Aggarwal CC, editor. Data classification: algorithms and applications. CRC press; 2014.
[39] Zeng X, Martinez TR. Distribution-balanced stratified cross-validation for accuracy estimation. J Exp Theor Artif Intell 2000;12(1):1–12.
[40] Pandas, 2020. https://pandas.pydata.org/. [Accessed 7 April 2020].
[41] Scikit-learn. Machine learning in Python; 2018, https://scikit-learn.org/stable/, . [Accessed 7 March 2020].
[42] StackingClassifier. An ensemble-learning meta-classifier for stacking; http://rasbt.github.io/mlxtend/user_guide/classifier/StackingClassifier/; 2019, . [Accessed 7 March 2020].
[43] Fortin FA, De Rainville FM, Gardner MA, Parizeau M, Gagne C. DEAP: evolutionary algorithms made easy. J Mach Learn Res 2012;13:2171–5.
[44] DEAP documentation; 2020, https://deap.readthedocs.io/en/master/, . [Accessed 7 March 2020]..
[45] Hammad M, Plawiak P, Wang K, Acharya UR. ResNet—attention model for human authentication using ECG signals. Expert Syst 2020;e12547.
[46] Tuncer T, Ertam F. Neighborhood component analysis and reliefF based survival recognition methods for Hepatocellular carcinoma. Phys A Stat Mech Appl 2020;540123143.
[47] Zhang ZM, Tan JX, Wang F, Dao FY, Zhang ZY, Lin H. Early diagnosis of hepatocellular carcinoma using

machine learning method. Front Bioeng Biotechnol 2020;8:254.
[48] Abdar M, Zomorodi-Moghadam M, Das R, Ting IH. Performance analysis of classification algorithms on early detection of liver disease. Expert Syst Appl 2017;67:239–51.
[49] Arian R, Hariri A, Mehridehnavi A, Fassihi A, Ghasemi F. Protein kinase inhibitors' classification using K-nearest neighbor algorithm. Comput Biol Chem 2020107269.
[50] Dureja H, Gupta S, Madan AK. Topological models for prediction of pharmacokinetic parameters of cephalosporins using random forest, decision tree and moving average analysis. Sci Pharm 2008;76(3):377–94.
[51] Perdana RS, Pinandito A. Combining likes-retweet analysis and naive bayes classifier within twitter for sentiment analysis. J Telecommun Electron Comput Eng (JTEC) 2018;10 (1–8):41–6.
[52] Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol (TIST) 2011;2(3):1–27.
[53] Tsai CF, Li ML, Lin WC. A class center based approach for missing value imputation. Knowledge Based Syst 2018;151:124–35.
[54] Walker SH, Duncan DB. Estimation of the probability of an event as a function of several independent variables. Biometrika 1967;54(1-2):167–79.
[55] Ibrahim W, Abadeh MS. Protein fold recognition using Deep Kernelized Extreme Learning Machine and linear discriminant analysis. Neural Comput Appl 2019;31 (8):4201–14.
[56] Plawiak P, Abdar M, Plawiak J, Makarenkov V, Acharya U. DGHNL: a new deep genetic hierarchical network of learners for prediction of credit scoring. Inf Sci (Ny) 2019;516:401–18.
[57] Acharya UR, Raghavendra U, Fujita H, Hagiwara Y, Koh JE, Hong TJ, et al. Automated characterization of fatty liver disease and cirrhosis using curvelet transform and entropy features extracted from ultrasound images. Comput Biol Med 2016;79:250–8.
[58] Acharya UR, Faust O, Molinari F, Sree SV, Junnarkar SP, Sudarshan V. Ultrasound-based tissue characterization and classification of fatty liver disease: a screening and diagnostic paradigm. Knowledge Based Syst 2015;75: 66–77.
[59] Ali L, Wajahat I, Golilarz NA, Keshtkar F, Bukhari SAC. LDA–GA–SVM: improved hepatocellular carcinoma prediction through dimensionality reduction and genetically optimized support vector machine. Neural Comput Appl 2020;1–10.