



OPEN

## An innovative ensemble model based on deep learning for predicting COVID-19 infection

Xiaoying Su<sup>1</sup>, Yanfeng Sun<sup>2</sup>, Hongxi Liu<sup>1</sup>, Qiuling Lang<sup>1</sup>, Yichen Zhang<sup>1</sup>, Jiquan Zhang<sup>3</sup>, Chaoyong Wang<sup>1✉</sup> & Yanan Chen<sup>1</sup>

Nowadays, global public health crises are occurring more frequently, and accurate prediction of these diseases can reduce the burden on the healthcare system. Taking COVID-19 as an example, accurate prediction of infection can assist experts in effectively allocating medical resources and diagnosing diseases. Currently, scholars worldwide use single model approaches or epidemiology models more often to predict the outbreak trend of COVID-19, resulting in poor prediction accuracy. Although a few studies have employed ensemble models, there is still room for improvement in their performance. In addition, there are only a few models that use the laboratory results of patients to predict COVID-19 infection. To address these issues, research efforts should focus on improving disease prediction performance and expanding the use of medical disease prediction models. In this paper, we propose an innovative deep learning model Whale Optimization Convolutional Neural Networks (CNN), Long-Short Term Memory (LSTM) and Artificial Neural Network (ANN) called WOCLSA which incorporates three models ANN, CNN and LSTM. The WOCLSA model utilizes the Whale Optimization Algorithm to optimize the neuron number, dropout and batch size parameters in the integrated model of ANN, CNN and LSTM, thereby finding the global optimal solution parameters. WOCLSA employs 18 patient indicators as predictors, and compares its results with three other ensemble deep learning models. All models were validated with train-test split approaches. We evaluate and compare our proposed model and other models using accuracy, F1 score, recall, AUC and precision metrics. Through many studies and tests, our results show that our prediction models can identify patients with COVID-19 infection at the AUC of 91%, 91%, and 93% respectively. Other prediction results achieve a respectable accuracy of 92.82%, 92.79%, and 91.66% respectively, f1-score of 93.41%, 92.79%, and 92.33% respectively, precision of 93.41%, 92.79%, and 92.33% respectively, recall of 93.41%, 92.79%, and 92.33% respectively. All of these exceed 91%, surpassing those of comparable models. The execution time of WOCLSA is also an advantage. Therefore, the WOCLSA ensemble model can be used to assist in verifying laboratory research results and predict and to judge various diseases in public health events.

Public health events have become increasingly common since globalization, with COVID-19 being the most prominent example. However, due to the multiple mutations that COVID-19 undergoes during transmission, it is crucial to identify infections safely and efficiently. Previous research on predicting COVID-19 has primarily focused on monolithic neural network models<sup>1,2</sup>. ANN provides direction in the medical decision-making process. Researchers used an artificial neural network model to study the status of patients with COVID-19<sup>3</sup>. In addition, X-ray images are used as input data, and the CNN model is used to classify and evaluate the severity level of COVID-19 patients<sup>4</sup>. A monolithic model has low prediction accuracy. Although a small number of studies have used mixed models to predict the development trend of COVID-19<sup>5-9</sup>. Qu et al.<sup>10</sup> proposed an innovative ensemble model that utilizes uses a new whale optimization algorithm (SCWOA) to predict COVID-19. This approach assigns different weights to each neural network model and optimizes the weight of the ensemble model. The results of the tests demonstrate that different deep learning models have varying prediction performance. Wang Gang compared the differences between CARIMA, es, GRNN and ARIMA-GRNN in predicting the number of daily cases. This result demonstrated that the ARIMA-GRNN mixed model<sup>8</sup> is better. Mohan et al.<sup>11</sup> used Supervised machine learning model (EAMA) for predicting long-term COVID-19 related parameters within India and globally. Kuvvetli et al.<sup>7</sup> predicted the number of daily deaths and cases of COVID-19

<sup>1</sup>School of Jilin Emergency Management, Changchun Institute of Technology, Changchun 130021, China. <sup>2</sup>College of Computer Science and Technology, Jilin University, Changchun 130012, China. <sup>3</sup>School of Environment, Northeast Normal University, Changchun 130024, China. ✉email: wangchaoyong@cit.edu.cn

with three different ANN models. To predict COVID-19 severity, logistic regression and ANN were applied to construct the COVID-19 severity prediction model<sup>12</sup>. Anum Shafiq et al.<sup>13</sup> predict mortality rate of COVID-19 in Italy. Moreover, there are many methods for establishing prediction models in the medical field, yet they are only used in other aspects of medicine, which is different from the direction of this article using clinical data to study COVID-19<sup>14–17</sup>, such as epilepsy<sup>18,19</sup>, Cholera<sup>20,21</sup>, Ebola<sup>22–25</sup>. Machine learning provides a good effect and a reliable explanation for the study of diseases. Deep Learning as a Data-Dependent alternative to predict the development of the US COVID-19 epidemic<sup>26</sup>. Although these models have made significant contributions to epidemic prevention and control, existing research rarely uses laboratory indicators of patients to predict COVID-19 infection in biomedicine, and their prediction performance is poor. In this respect, in order to provide an accurate prediction model for medical institutions, we put forward an innovative deep learning model.

Optimizing and adjusting parameters using correlations between them is a critical step towards achieving optimal solutions for deep learning models<sup>27–30</sup>. Usually, the parameters are determined through a trial and error approach which is often fraught with errors and particularities, making it difficult to predict the approximate optimal solution in most cases, thus affecting the prediction performance of model. The Convolutional Neural Networks, Long-Short Term Memory (CNNLSTM) model<sup>31</sup>, which combines both spatial and temporal depth, holds significant promise. Brodzicki et al.<sup>29</sup> has investigated the function of the whale optimization algorithm with a simple mechanism, few parameters, and strong optimization ability, and simulated the three-dimensional characteristics of the original Whale Optimization Algorithm (WOA). The results demonstrate that WOA is an excellent algorithm, that can optimize parameters more effectively than other well-known heuristic optimization algorithms.

This paper presents a new system for detecting COVID-19 by enhancing the original deep learning model. Based on the research of Alakus and Turkoglu<sup>31</sup>, we employ the Whale Optimization Algorithm to reconstruct the architecture of the CNNLSTM model and optimize its parameters. Therefore, a new ensemble model is proposed to predict COVID-19 infection. In this paper, key contributions can be summarized as the following:

- Optimization of the internal parameters of the CNNLSTM using WOA, and reconstructing its architecture. In addition, we use Grey Wolf Optimizer (GWO) to optimize the internal parameters of CNNLSTM, and take it as one of the comparative experiments of different deep learning models (GWOLSA).
- The ensemble model (WOCLSA) is proposed to improve the prediction performance.
- The innovative ensemble model can be used to predict various public health diseases and provide help for medical disease diagnosis.

The structure of this study is as follows. Section "Methods" describes three neural network approaches, generalization of WOA and generalization of GWO. Section "Innovating the framework of forecasting model" describes three hybrid models as well as the WOCLSA. Section "Data description and evaluation criteria" summarizes the data sources and specific performance evaluation standards. Section "Experimental results and analysis" talks about the prediction results of the proposed model and compares the results with those of other models. Section "Discussion" is the discussion of this paper. At last, Section "Conclusion" is the conclusion of this paper.

## Methods

**CNN.** CNN is a deep neural network model with convolutional computation. It adds the parameters of the convolution kernel to extract the features of the data. The parameters of the convolutional layer include three parts: size, step size, and padding.

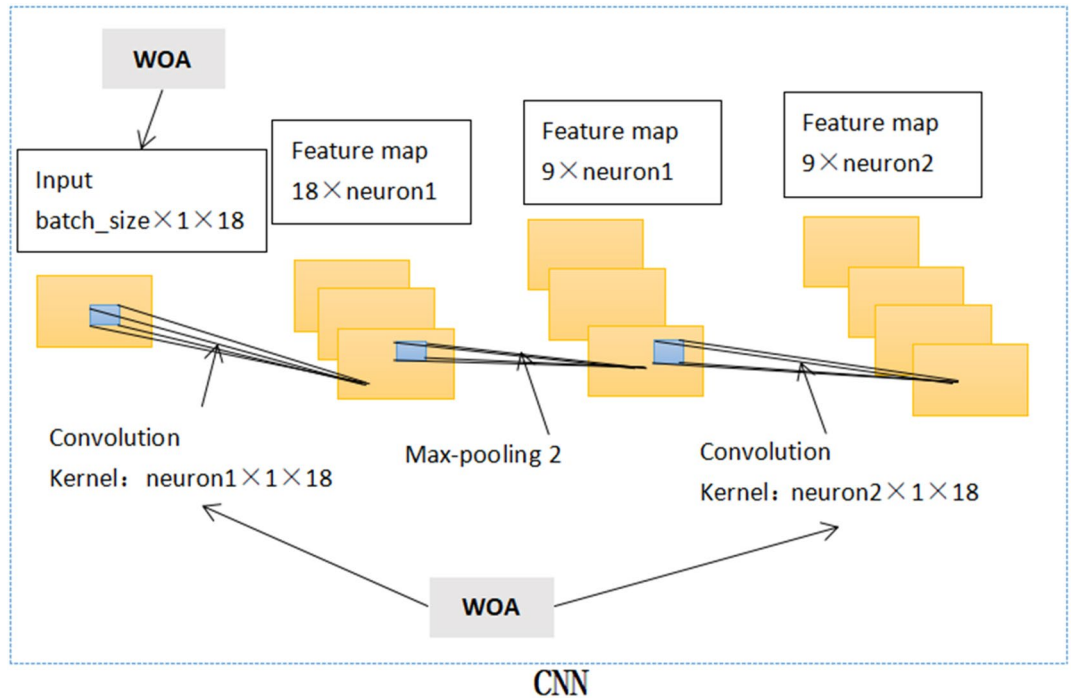
Undoubtedly, these parameters above are crucial. They jointly determine the output characteristics of CNN. In this paper, the training data from 480 patients was processed by batch size and entered into the deep learning model in batches. The amount of batch size is determined by WOA. Figure 1 shows the CNN internal structure of the WOCLSA deep learning model. CNN does the first convolution to generate the feature map. The number of filters for the first convolution kernel is set to neuron 1. Feature maps are pooled through the pooling layer. Next, CNN does the second convolution. The number of filters for the second convolution kernel is set to neuron 2. Batch size, neuron 1 and neuron 2 use WOA to find the optimal parameter values. For the other two deep learning models, fixed parameters are adopted. The number of filters for the first convolution kernel is set to 500 and the number of filters for the second convolution kernel is set to 250. Batch size is set to the default parameter. Figure 2 shows the CNN internal structure of Convolutional Neural Networks (CNN), Long-Short Term Memory (LSTM) and Artificial Neural Network (ANN) (CNNLSTMA) and CNNRNN deep learning models. Figure 3 shows the CNN internal structure of the GWOLSA.

**LSTM.** LSTM is a time recurrent neural network, which is an advanced version of Recurrent Neural Networks (RNN). LSTM has a more complex structure and has advantages in long sequence problems. It is divided into the following four departments:

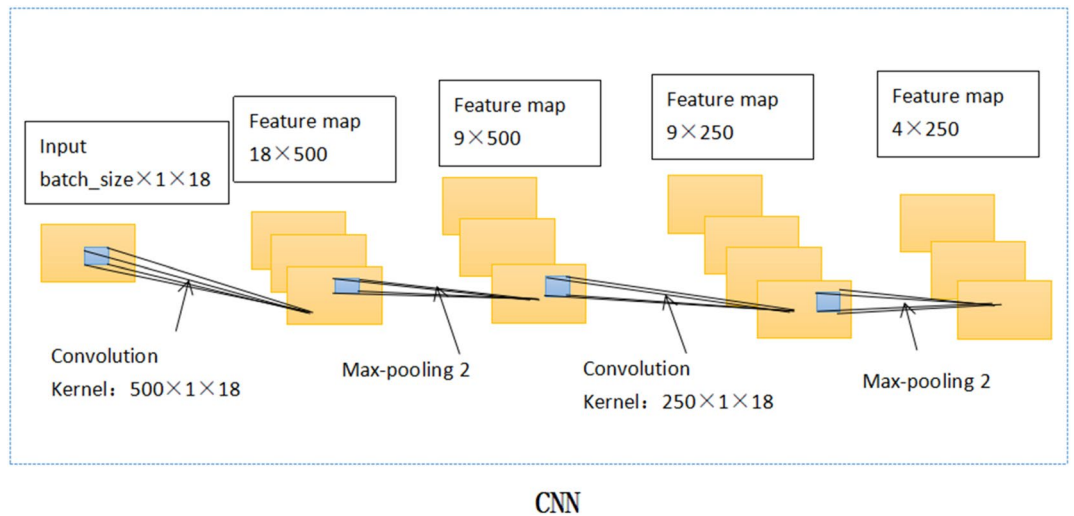
First, LSTM uses the forgetting gate to determine the information carried by the cells on the upper layer.  $x_t$  and  $h_{t-1}$  are input and activated by the sigmoid function to obtain  $f$ . The output value ranges from 0 to 1.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

Then there is the update gate, which is divided into two parts. One is that this part can be seen as accepting new information, and the tanh activation function normalizes the content in  $[-1,1]$ . The other is that it can be seen as the content of new information retention.



**Figure 1.** CNN internal structure of the WOCLSA.



**Figure 2.** CNN internal structure of CNNLSTMA and CNNRNNA models.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

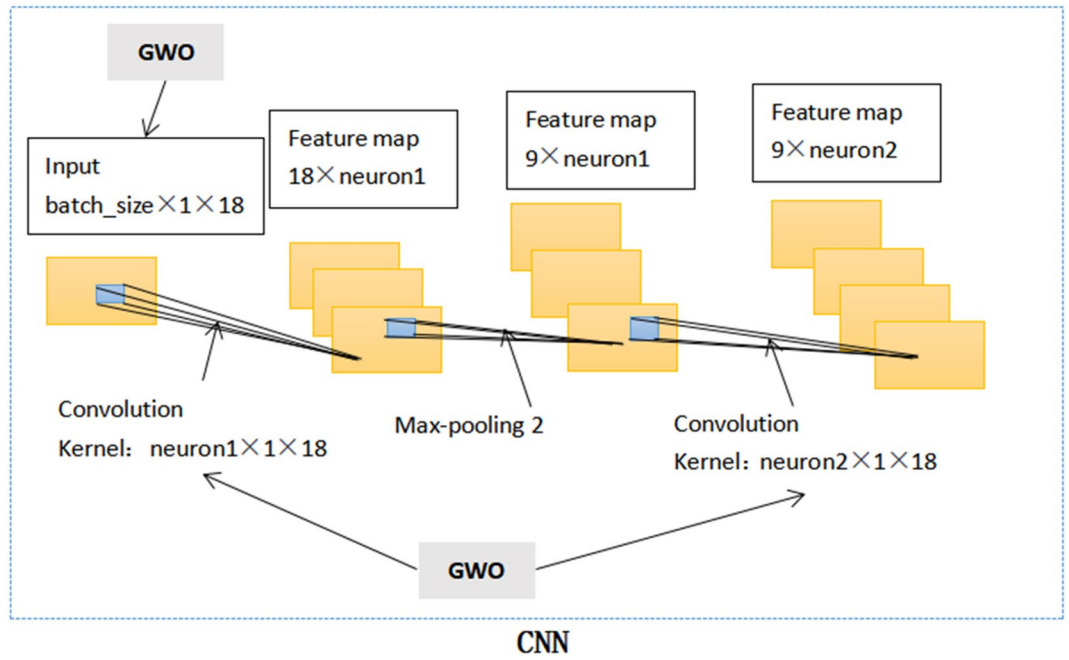
$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

The following operation is to update  $c_t$  and selectively forget the previous information to become a new  $c_t$ .

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

Finally, it is the output of LSTM. At this time, the neuron  $c_t$  has been updated, using the ReLU activation function to output the content. The  $c_t$  is scaled by tanh and multiplied by  $O_t$ . Next, put the cell state  $h_t$ .

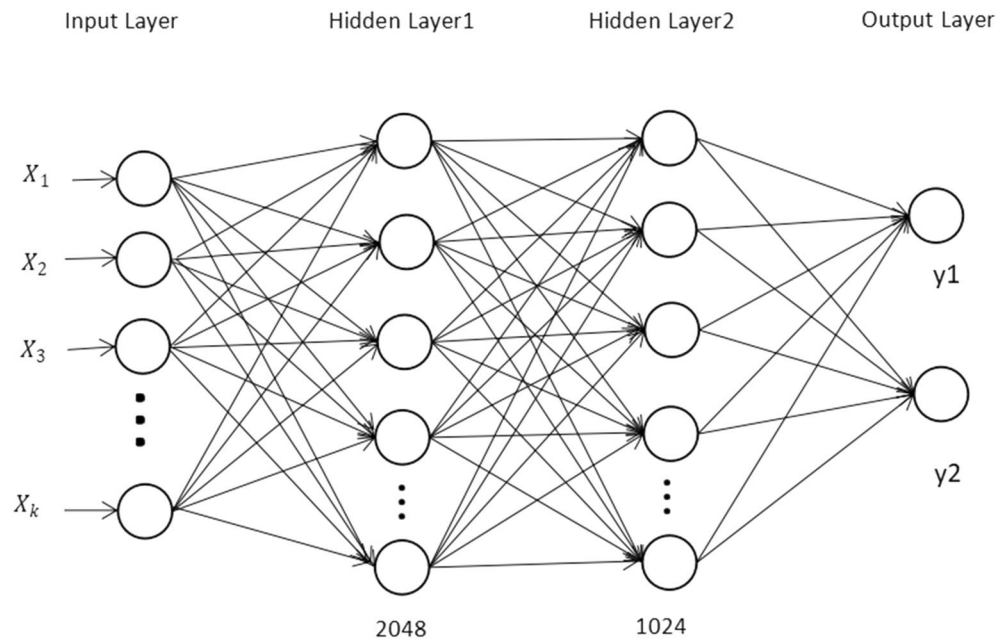
$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{5}$$



**Figure 3.** CNN internal structure of the GWOCLSA deep learning model.

$$h_t = O_t * \tanh(C_t) \tag{6}$$

**ANN.** ANN is a deep learning model that imitates the structure of the brain’s neural network. It contains a large number of neurons. When the input data enters the model, the weights of the neural network are updated by calculating the gradient of the error function through the back propagation algorithm, so that the output of the neural network is closer to the expected value. In this study, Fig. 4 is the structure of ANN. In the WOCLSA model, when the data is output from LSTM, the dimension of the output space is 512. The number of data entering the input layer is determined by the batch size. Figure 4 shows that there are k data sequences. We set the neuron parameters of hidden layer1 to 2048, and hidden layer2 to 1024 respectively.



**Figure 4.** Basic structure of ANN.

**Generalization of WOA.** Mirjalili and Lewis<sup>32</sup> proposed the WOA. The WOA is a heuristic intelligent optimization algorithm that mimics the behavior of whale predation. If a whale finds prey, then other whales will surely swim to the whale that finds prey to compete for prey. In this paper, one solution represents a whale, and several solutions are represented by several whale individuals. In the process of using WOA to search for the solution of the problem, it can be seen that several whale individuals are constantly updating their individual positions until a satisfactory solution is found. The main behaviors include three parts: encircling prey, bubble-net attacking, and searching for predation. The following is the mathematical model of different stages of the WOA process.

*Encircling prey.* When the whale is encircling the prey, the whale position of the search agent will increase the number of iterations needed to update the position vector. These position vectors are related to the vector position of the best whale individual. We use a four-dimensional whale search dimension and 10 whale counts, setting the upper and lower limits of the search. In order to prevent the required parameters of neuron 1, neuron 2, dropout and batch size from overflowing, an adaptation function is set to correct these parameters.

$$Z = |Q \cdot W^*(t) - W(t)| \quad (7)$$

$$W(t+1) = W^*(t) - C \cdot Z \quad (8)$$

where  $t$  represents the current iteration,  $C$  and  $Q$  represent the coefficient vectors,  $W^*$  is the position vector of the optimal solution obtained in this way, and  $W$  represents the position vector of the whale. The values of  $C$  and  $Q$  are calculated according to the formula:

$$C = 2a \cdot r - a \quad (9)$$

$$Q = 2r \quad (10)$$

As the number of iterations increases,  $a$  decreases linearly from 2 to 0, and  $r$  is a random vector in  $[0,1]$ .

*Bubble-net attacking.* Bubble net attack is the bubble predation behavior of whales. The attack of the bubble net is divided into contraction encirclement and spiral update position mechanisms. In order to realize the contraction surround behavior, it is almost the same as the mathematical model of the prey behavior mentioned above. The difference is that the value of  $C$  changes from  $[-a, a]$  to  $[-1,1]$ . In the stage of spiral update position mechanism, the search agent approaches the target in a spiral shape. For the next iteration, probability selection will be used. If the probability is less than 0.5, the shrink-wrap mechanism is adopted. The probability is greater than 0.5, and the spiral update position is used.

$$Z' = |W^*(t) - W(t)| \quad (11)$$

$$W(t+1) = Z' \cdot e^{bk} \cdot \cos 2\pi k + W^*(t) \quad (12)$$

where  $W^*$  is the position vector of the best search agent,  $W(t)$  is the position vector of the search agent in iteration  $t$ ,  $Z'$  is the coefficient vector.  $b$  is a constant value. Random numbers are represented by  $k$ , and the range is  $[-1,1]$ .

*Search for prey.* At this stage, if the predation of  $C$  does not belong to  $[-1,1]$ , then the search agent whale randomly selects one from the current whale population instead of being close to the best whale agent. In order to avoid falling into the state of local optimal solution. The mathematical model is as follows:

$$Z = |Q * W_{rand} - W(t)| \quad (13)$$

$$W(t+1) = W_{rand} - C \cdot Z \quad (14)$$

$W_{rand}$  is the randomly selected location of the whale,  $W(t)$  is the position of the search agent after  $t$  iterations. This paper has a total of three iterations.  $C$  and  $Q$  are position vectors, and  $Q$  is the coefficient vector<sup>33</sup>.

**Generalization of GWO.** The Grey Wolf Optimizer (GWO) is also a swarm intelligence algorithm. The algorithm is an optimization search method developed to mimic the prey predation of gray wolves. The GWO process consists of three main steps: encircling, hunting, and attacking the prey<sup>34</sup>. The whale optimization algorithm is similar to the gray wolf optimization algorithm. Comparing with the whale optimization algorithm, we also set up 10 Gy wolves and had a total of three iterations. The range of prey searched by the gray wolf is also set to 4-dimensional.

**WOA algorithm optimizes deep learning model.** Deep learning methods generally involve the following stages:

1. Data pre-processing.
2. CNNLSTM architecture update and reconstruction.
3. Performing model learning and data fitting.

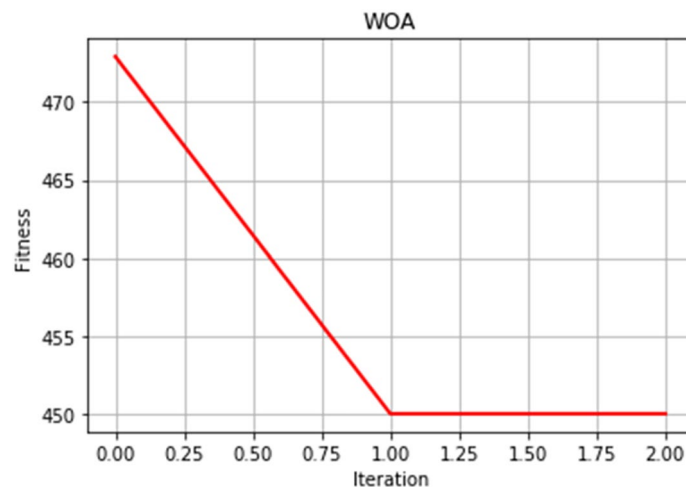
The last stage is crucial for parameter selection. In this paper, we optimize some parameters using WOA such as dropout or batch size<sup>10</sup>.

In this paper, the number of training samples is small. In order to prevent the overfitting phenomenon which leads to distortion of prediction results, we apply the dropout function, which will enhance the generalization of the model. To explore the impact of the dropout function on the prediction accuracy and mitigate the risk of overfitting, this study employs the whale algorithm to optimize the parameters of the dropout function. This approach aims to identify an appropriate initial value that can lead to more consistent results. In addition, the size of batch size affects the reliability and efficiency of the model. It is not appropriate for the batch size to be too large or too small. Setting a reasonable batch size in the model can improve the training speed and make the gradient's descending direction more accurate. Therefore, this model also optimizes the batch size parameter.

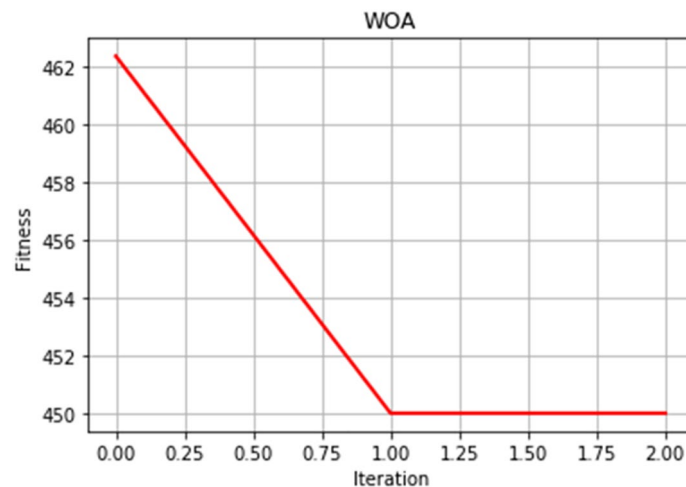
The convergence speed of the optimal fitness curve is a key indicator to evaluate the excellence of the WOA. The smoother and better the optimal fitness curve, the faster can converge to the optimal solution quickly. We adjust a four-dimensional whale search dimension. This experiment sets the whale search range, iteration times, and population number. After many comparative experiments, a target value close to the actual optimal value of the function is found. Figures 5 and 6 selected two experimental results.

### Innovating the framework of forecasting model

**Three ensemble models.** In general, some models such as CNN and Recurrent Neural Networks (RNN) are used to predict medical diseases<sup>35</sup>. RNN is a kind of recursive neural network that uses sequence data as input. RNN solves the input problem through its internal memory, and all using the same functions in the process. The integrated model Artificial Neural Network (ANN), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) is called CNNRNNNA, and the data results are continued into the RNN model to run after the CNN completes its convolution<sup>36,37</sup>. Then, the LSTM dimension of the output space enters



**Figure 5.** Optimal fitness curve diagram.



**Figure 6.** Optimal fitness curve diagram.

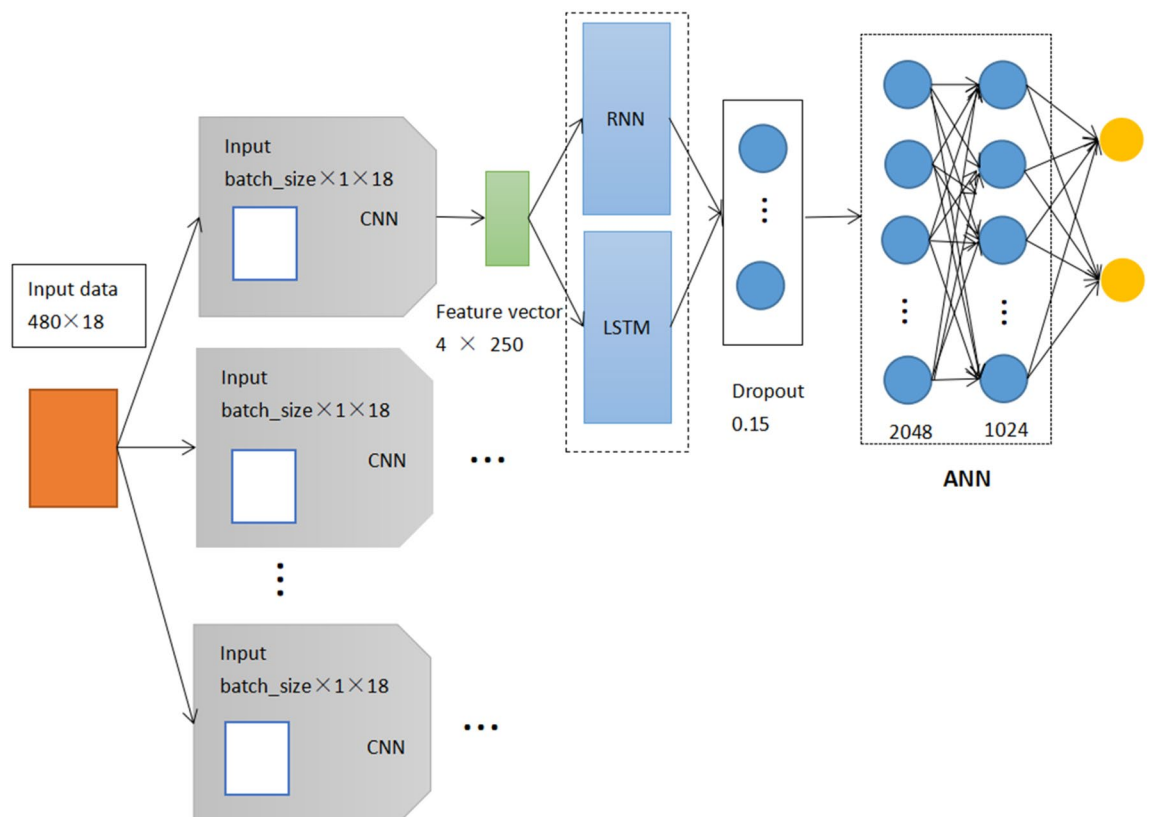
into neural network learning. Compared with classical RNN, LSTM can alleviate the gradient disappearance or explosion phenomenon of RNN. In addition, CNNLSTM has shown good prediction results due to this integrated model in Alakus and Turkoglu's<sup>31</sup> prediction, CNNLSTM combines the respective spatial and temporal advantages of ANN, CNN and LSTM, with the flexibility and efficiency of application in predicting public health diseases<sup>38–42</sup>. In this paper, the mixed model is obtained by adjusting the parameters of CNNLSTM model, which is used as a comparative prediction of the WOCLSA model, and the mixed model is called CNNLSTMA which includes CNN, LSTM and ANN. Therefore, the effects of CNNRNN and CNNLSTMA mixed models are worth studying. Therefore, two hybrid models were used as comparative study experiments, Fig. 7 shows the basic structure of the CNNRNN and CNNLSTMA.

We use the Gray Wolf Optimization Algorithm to optimize the neuron number, dropout and batch size parameters of CNNLSTM and change the model structure of CNNLSTM. It is referred to as the GWOCLSA deep learning model. In order to better verify the prediction performance of the WOCLSA model, we added the GWOCLSA deep learning model as a comparison of heuristic optimization algorithms. Figure 8 shows the basic structure of GWOCLSA.

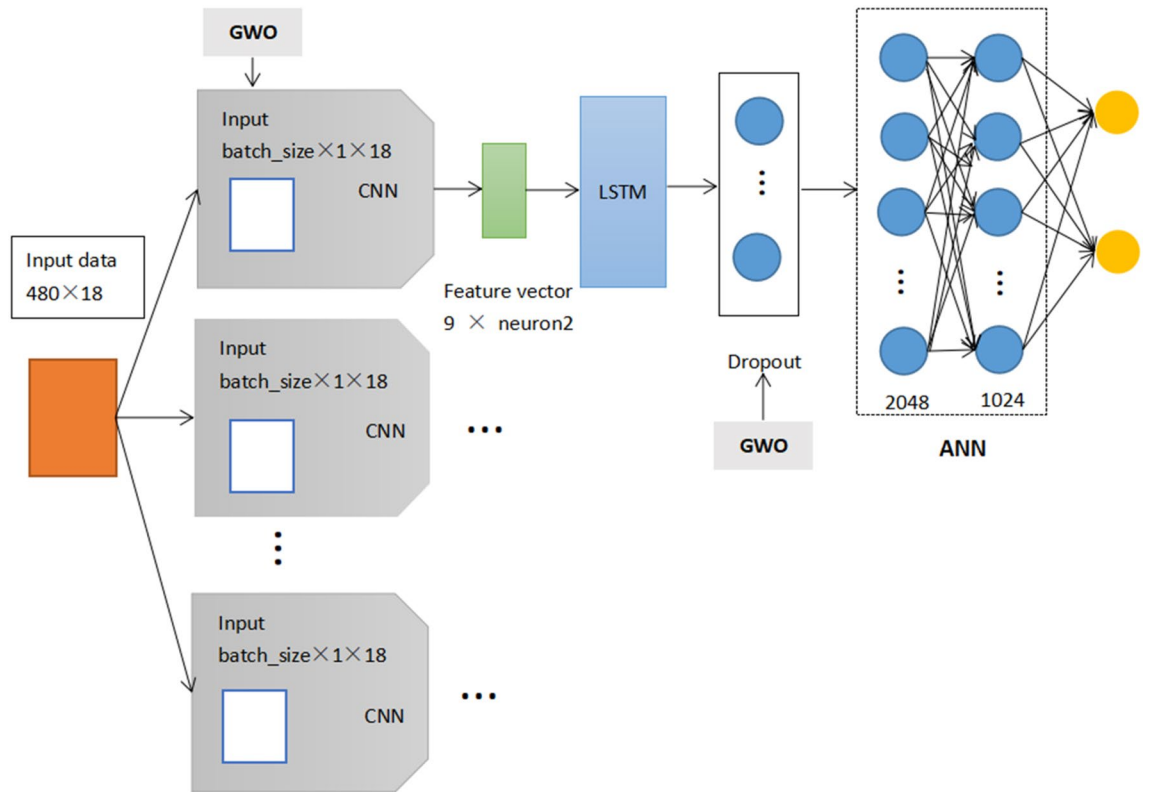
**WOCLSA.** In this paper, the patient's data is input into four different deep learning models respectively, predicting the performance of these models and analyzing the results with five different evaluation metrics. Figure 9 images the flowchart of this paper.

In the WOCLSA deep learning model, the patient's data is fed into the model that ANN, CNN, and LSTM monolithic models are architecturally integrated. After the whale algorithm optimizes parameters of the integrated model, the output values are available in the classifier after the activation function to get the prediction results. Figure 10 shows the framework of the WOCLSA prediction model.

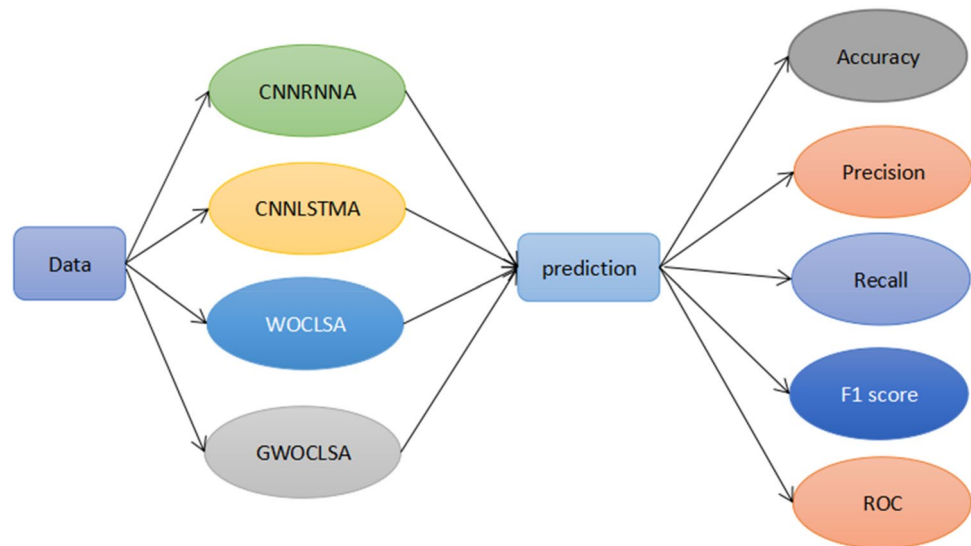
Based on Alakus and Turkoglu's<sup>31</sup> study, the parameters in the original model were set for each model using the trial and error approach, which has certain errors and lacks universality. In this study, the CNNLSTM architecture was rebuilt and the parameters of dropout and the batch in the CNNLSTM model were optimized by WOA algorithm to find the global optimal solution parameters. When the complete data set is run through the model once and the results are returned once, this process completes an epoch. For all the integrated models, we set 100 epochs. It can make the model performance best and reduce over-fitting. Figure 11 shows the detailed design of WOCLSA.



**Figure 7.** The basic structure of the CNNRNN and CNNLSTMA.



**Figure 8.** The basic structure of the GWOLSA.

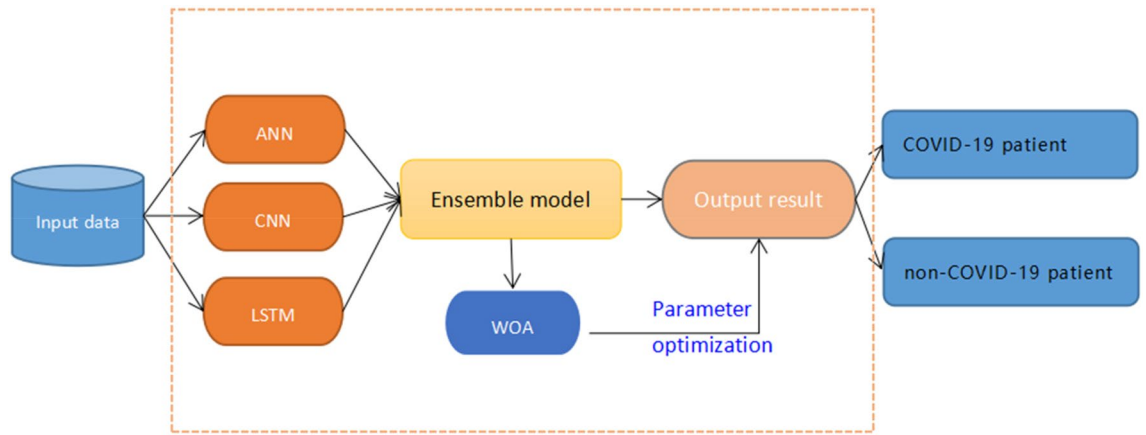


**Figure 9.** Flowchart of this study.

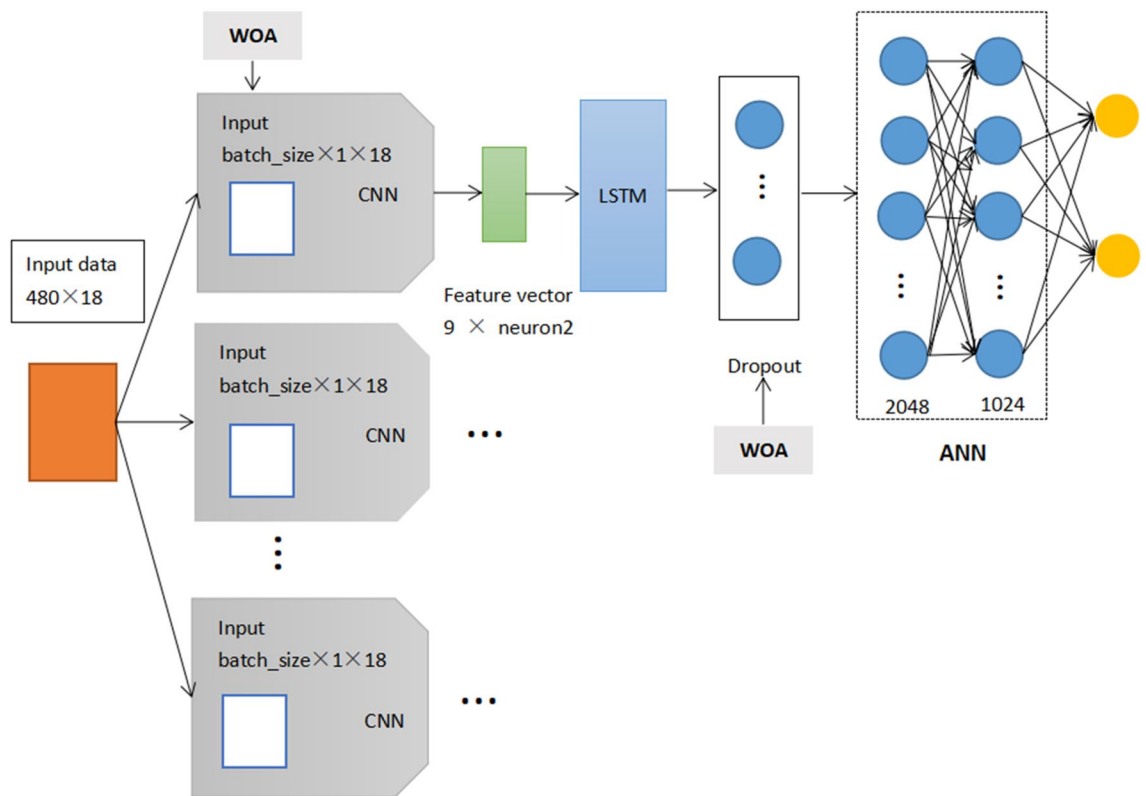
**Data description and evaluation criteria**

**Data description.** The dataset includes the laboratory findings of the patients seen at the Hospital Israelita Albert Einstein at Sao Paulo Brazil<sup>43</sup>. In the above data set, there are 5644 various patients with 111 laboratory findings. There are 10% positive patients in the data, where 18 indicators in the findings were important for COVID-19 disease, as shown in Table 1 below, and after the balancing process, there are 18 laboratory findings from 600 patients left<sup>31</sup>. In the balanced dataset, we have 520 non-COVID-19 patients and 80 COVID-19 patients. Our experiment divides 600 patients into test and training subsets, where the training subset contains 480 patients and the test subset contains 120 patients. In cases of relatively small samples, k fold cross-validation approach and train-test split approach are used in health studies. Based on Comparison of deep learning approaches to predict COVID-19 infection the paper, for the hybrid model, tenfold cross validation prediction





**Figure 10.** The framework of the WOCLSA prediction model.



**Figure 11.** Detailed design diagram of WOCLSA.

Patient indicators	Hematocrit, hemoglobin, platelets, red blood cells, lymphocytes, leukocytes, basophils, eosinophils, monocytes, serum glucose, neutrophils, urea, C reactive protein, creatinine, potassium, sodium, alanine transaminase, aspartate transaminase
--------------------	---

**Table 1.** Patient indicators.

performance results are slightly lower than the train-test split method<sup>31</sup>. Therefore, we use 80–20 train-test split approach to validate the data. Table 1 below shows the indicators.

**Evaluation standards.** We predict the performance of the model through five methods. They are accuracy, f1 score, accuracy, recall, and area under the ROC curve (AUC)<sup>31,44</sup> respectively. In this section, we will summarize each evaluation standard as follows:

Accuracy ( $A$ ) refers to the number of correctly classified samples as a percentage of the total number of samples.  $N$  represents the sample of all predicted COVID-19.

$$A = (tp + tn)/N \quad (15)$$

$$N = tp + tn + fp + fn \quad (16)$$

In the above statement,  $tp$  represents true positive,  $tn$  represents true negative,  $fp$  represents false positive, and  $fn$  represents false negative. Accuracy is a key research characteristic that measures the proximity of sample parameters to the overall characteristics<sup>45</sup>. By demonstrating the accuracy of a study through appropriate metrics, researchers can establish the reliability of their findings. Higher accuracy values signify greater reliability.

Precision can be defined as the ratio of correctly predicted positive observations to the total predicted positive observations. It is a performance metric in neural networks.

$$P = tp/(tp + fp) \quad (17)$$

Recall refers to the ratio of predicted correct positive cases to actual positive cases.

$$R = tp/(tp + fn) \quad (18)$$

In this paper, recall rate is the ratio of correctly identified COVID-19 patients to the total number of COVID-19 patients. The closer the recall score must be to 1 in order to be perfectly classified.

F1 score is the weighted average of precision and recall values. In other words, The higher the F1 score, the better the prediction effect.

$$F1 = 2/(1/P + 1/R) = 2 * P * R / (P + R) \quad (19)$$

Finally, the AUC score has an important role in medical research. AUC is a measure of the classifier's ability to distinguish between positive and negative classes. In order to predict the disease better, the closer the above evaluation criteria are to 1.

## Experimental results and analysis

**Experimental equipment.** During all the experiments, we use the Spyder (3.2.3) development tool and employ the Python 3.6 programming language to execute in Tensorflow (1.2.1) exploitation environment. Our development code is based on the program provided by the author of Alakus and Turkoglu<sup>31</sup>.

**Results and analysis.** In order to compare the time execution efficiency of the WOCLSA and GWOCLSA, we select the results of three experiments to study. The following conclusions can be drawn from Table 2, the execution time of WOCLSA is less than GWOCLSA every time. Therefore, WOCLSA is relatively efficient.

We have performed many test experiments, the above are the results of three tests. From Table 3, we can see that after the parameter adjustment of the whale optimization algorithm, the area under the ROC curve has been significantly improved. At the same time, the prediction rate has been significantly enhanced, with the accuracy, F1 score, precision rate and recall value all exceeding 92%. The experimental value of WOCLSA exceeds that

	Model	Total time
First time	WOCLSA	<b>94 min</b>
	GWOCLSA	108.9 min
Second time	WOCLSA	<b>98.2 min</b>
	GWOCLSA	121.6 min
Third time	WOCLSA	<b>99.8 min</b>
	GWOCLSA	109.6 min

**Table 2.** The execution time of two deep learning model. Significant values are in bold.

	Accuracy	F1-score	Precision	Recall	AUC
CNNRNNNA	0.8655	0.8655	0.8655	0.8655	0.74
CNNLSTMA	0.8657	0.8742	0.8742	0.8742	0.83
WOCLSA	<b>0.9282</b>	<b>0.9341</b>	<b>0.9341</b>	<b>0.9341</b>	<b>0.91</b>
GWOCLSA	0.9003	0.9160	0.9160	0.9160	0.93

**Table 3.** Evaluation standards of all ensemble models with train-test split approach in test 1. Significant values are in bold.

of other hybrid models. Although AUC of GWOLSA is 93%, it is slightly higher than WOCLSA, the other evaluation indicators for GWOLSA are below WOCLSA. From Table 4, the accuracy, F1 score, precision rate and recall value of the WOCLSA model is respectively 92.79%, 92.79%, 92.79% and 92.79%. In the Table 5, these evaluation standards are still the highest compared to other models. The AUC values of the GWOLSA and WOCLSA models are close based on the results of multiple tests.

From Fig. 12, an AUC score over 90% is considered outstanding. WOCLSA achieved good results, it can be seen that it reached 91%. The AUC of GWOLSA is 93%, while the other two hybrid models by AUC scores were less than 90%. They do not play a good prediction effect. Figure 13 shows the AUC values for multiple tests of the WOCLSA model. The AUC has a high score of 91%, 95%, and 93%. The AUC of the WOCLSA and GWOLSA models are higher than that of CNNLSTM model researched by Alakus and Turkoglu<sup>31</sup>, which is 90%. Therefore, this is an innovative breakthrough in predictive performance. There will be some other factors in the process of prediction, such as the configuration of the computer and the nature of the errors in the calculated scores during the optimization of the heuristic algorithm. Although each prediction result is subject to small changes by these external factors, the overall high performance of the prediction is constant. The value of AUC plays a key role in medical disease classification<sup>45</sup>, a higher AUC score makes a more accurate prediction for real patients.

## Discussion

It is vital to use artificial intelligence systems to predict the clinical course of public health diseases. Various intelligent prediction models have been applied in numerous fields, providing medical professionals with valuable insights into clinical diagnostics. The country's prediction of public health disease events has always attracted much attention. At present, the simulation research on the clinical prediction of COVID-19 is very limited.

For example, Zongxi et al.<sup>46</sup> applied Gray Wolf optimization and multi-machine learning to predict COVID-19. The optimal weights of the ANFIS, LSSVM and LSTM models are searched by the Grey Wolf optimization algorithm, and the prediction model is established. Although machine learning is used in most studies of prediction in medicine, the predictive performance of deep learning still deserves to be explored and applied. The prediction accuracy of deep learning model has been improved, but it has been used to predict the epidemic trend in COVID-19, rather than clinical diagnosis. There are some models such as IPSO–DNN and MPSO–CNN that optimize parameters with intelligent algorithms<sup>47,48</sup>. Sun et al.<sup>49</sup> used a Genetic algorithm to optimize ANN. However, these models are relatively backward, and the prediction accuracy is low. We also have different evaluation criteria for predicting diseases from these models. In addition, Qian et al.<sup>50</sup> and Pasaribu et al.<sup>51</sup> respectively applied the Infectious disease dynamics model (SEIR) and the big data Model (GSTAR).

From Tables 3, 4, and 5, it can be seen that the prediction performance of the deep learning model optimized by whale algorithm has high reliability. As an optimization comparison of heuristic algorithm, the prediction performance of grey wolf algorithm is lower than that of whale algorithm.

At present, the lack of literature research includes the following three factors: there are few models for medical diagnosis based on patient laboratory indicators; most predict outbreak trends of COVID-19, rarely predict patient-specific disease; the prediction performance of existing deep learning models is low, and stable and efficient models are urgently needed.

Based on the above analysis, we propose an innovative ensemble model based on deep learning. The main contributions are as follows:

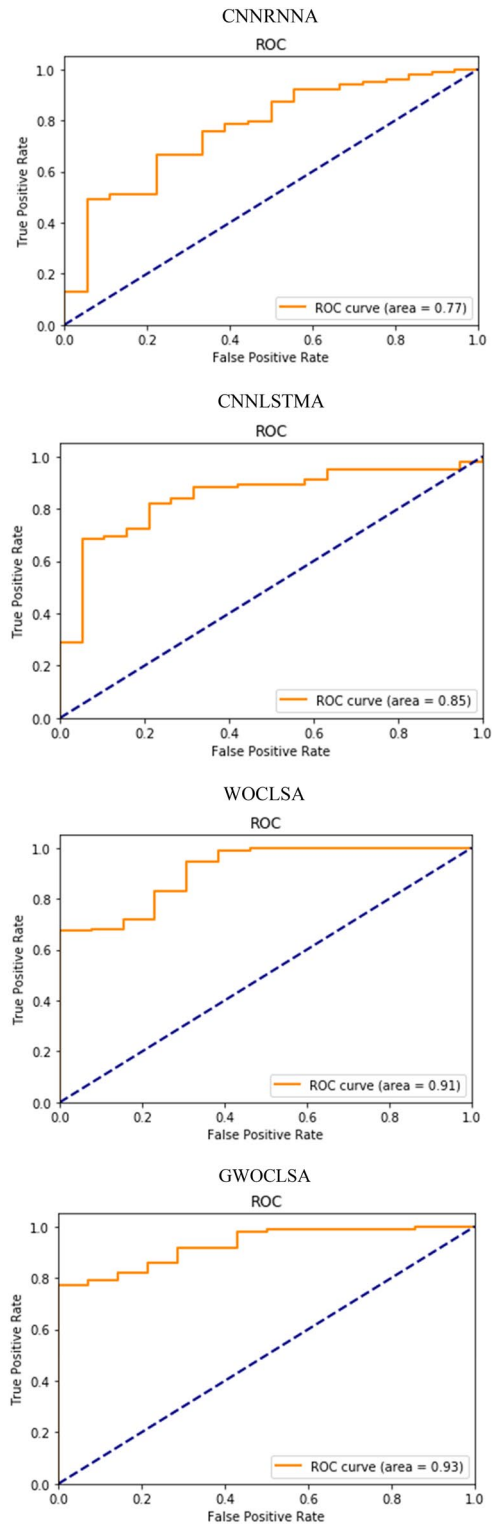
- Optimization of the internal parameters of the CNNLSTM using WOA, and reconstructing its architecture. As a heuristic algorithm, WOA is well combined with deep learning model. It has a better optimization effect than other heuristic algorithms.

	Accuracy	F1-score	Precision	Recall	AUC
CNNRNNNA	0.8655	0.8655	0.8655	0.8655	0.74
CNNLSTMA	0.8789	0.8788	0.8788	0.8788	0.85
WOCLSA	<b>0.9279</b>	<b>0.9279</b>	<b>0.9279</b>	<b>0.9279</b>	<b>0.91</b>
GWOLSA	0.9191	0.9211	0.9211	0.9211	0.90

**Table 4.** Evaluation standards of all ensemble models with train-test split approach in test 2.

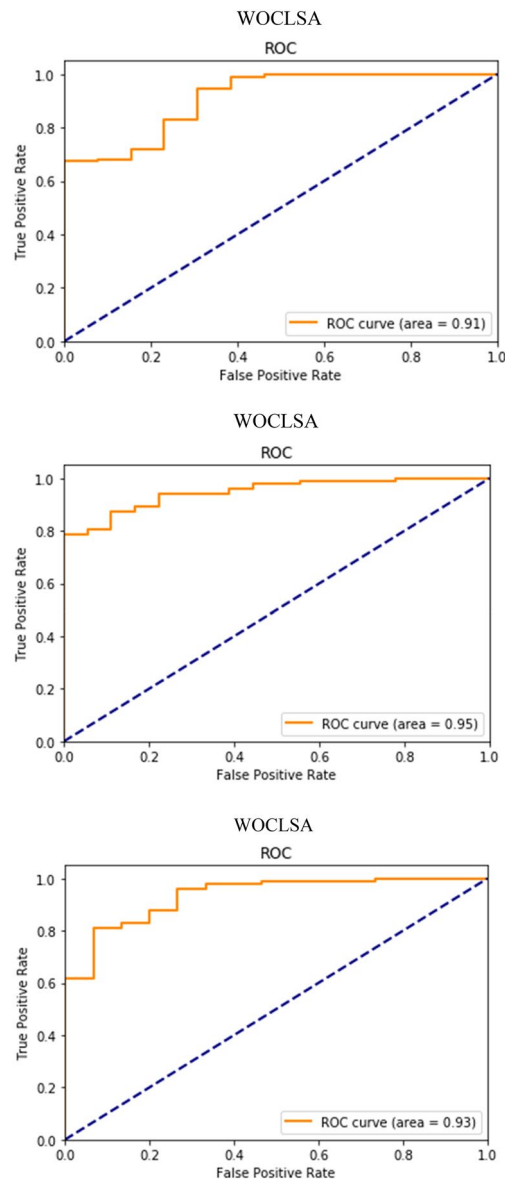
	Accuracy	F1-score	Precision	Recall	AUC
CNNRNNNA	0.8585	0.8535	0.8535	0.8535	0.79
CNNLSTMA	0.8890	0.8901	0.8901	0.8901	0.89
WOCLSA	<b>0.9166</b>	<b>0.9233</b>	<b>0.9233</b>	<b>0.9233</b>	<b>0.93</b>
GWOLSA	0.9099	0.9066	0.9066	0.9066	0.94

**Table 5.** Evaluation standards of all ensemble models with train-test split approach in test 3. Significant values are in bold.



**Figure 12.** AUC values of all models.

- The ensemble model (WOCLSA) is proposed to improve the prediction performance. The evaluation criterion of the model are superior to those of other deep learning models. The AUC is 91%, which is higher than previous research results.
- The innovative ensemble model can be used to predict various public health diseases and provide help for medical systems and laboratory research. Meanwhile, WOCLSA model prediction method can be used for in other fields to study and explore.



**Figure 13.** AUC values for multiple tests of the WOCLSA model.

## Conclusion

In the field of medical research, the accurate prediction of patients' diseases is a crucial component that enables effective clinical judgement and facilitates the improvement of treatment efficiency. A precise prediction can also help to alleviate the burden on hospitals and reduce the unnecessary waste of medical resources, which are often limited due to regional disparities in development. By prioritizing scarce medical resources, clinicians can ensure that essential epidemic prevention materials and emergency management strategies are deployed efficiently, thereby minimizing wastage and maximizing the benefits of available resources.

At present, most diseases are diagnosed by image recognition, and few studies are made by using patients' laboratory results. Research on COVID-19, most of them study the outbreak trend, and there is almost no research focused on personal laboratory indicators for medical diagnosis or laboratory research. Based on the research conclusion of Alakus and Turkoglu<sup>31</sup>, it is concluded that the prediction effect of mixed model is higher than that of single model, so we use different ensemble models for comparative forecasting. We compare the innovative ensemble model WOCLSA with the other three mixed models, and the performance of predicting COVID-19 infection is improved. It is a deep learning model lacking at present.

We employ a set of 600 data points for prediction. The results show the WOCLSA deep learning model outperforms other models in terms of prediction performance, with all evaluation criteria exceeding 90%. In terms of time efficiency, the program execution time of WOCLSA is less than that of GWOCLSA. Therefore, the WOCLSA model is able to enhance prediction performance and reduce the burden of doctors. In addition, the innovative ensemble model can reduce the error of manual diagnosis, provide a new public health disease model for prediction and play an important role in biomedical research.

For future research, the following directions can be pursued:

- Further studies could increase the sample of hospital patients and not be limited to the 600 laboratory study results. In this paper, the sample data used for prediction is relatively small. For the WOCLSA, the larger the dataset, the better it is to validate its predictive performance.
- Future studies should focus on optimizing index selection. This paper only considered 18 indicators, but medical diagnosis can benefit from additional or fewer indicators in the future.

## Data availability

The balanced data set of laboratory findings is derived from the study of Alakus and Turkoglu. The balanced dataset via <https://github.com/burakalakuss/COVID-19-Clinical>.

Received: 26 April 2023; Accepted: 25 July 2023

Published online: 29 July 2023

## References

1. Das, S. & Kolya, A. K. Predicting the pandemic: sentiment evaluation and predictive analysis from large-scale tweets on Covid-19 by deep convolutional neural network. *Evol. Intel.* **15**(3), 1913–1934 (2022).
2. Khan R U, Almakdi S, Alshehri M, et al. Probabilistic approach to COVID-19 data analysis and forecasting future outbreaks using a multi-layer perceptron neural network. *Diagnostics.* **12**(10), 2539 (2022).
3. Rathee, G. et al. ANN assisted-IoT enabled COVID-19 patient monitoring. *IEEE Access.* **9**, 42483–42492 (2021).
4. Irmak, E. COVID-19 disease severity assessment using CNN model. *IET Image Proc.* **15**(8), 1814–1824 (2021).
5. Lin, T. K. et al. Prediction of epidemic trend of COVID-19 in Western Guangdong and other prefecture-level cities: a study based on Holt two-parameter exponential smoothing model. *Pract. J. Cardiac Cerebral Pneumal Vasc. Dis.* **28**(2), 13–17 (2020).
6. Yang, J. et al. The prediction of the epidemic trend of COVID-19 using neural networks. *Ifac-papersonline.* **53**(5), 857–862 (2020).
7. Kuvvetli, Y. et al. A predictive analytics model for COVID-19 pandemic using artificial neural networks. *Decis. Anal. J.* **1**, 100007 (2021).
8. Wang, G. et al. Comparison of ARIMA, ES, GRNN and ARIMA–GRNN hybrid models to forecast the second wave of COVID-19 in India and the United States. *Epidemiol. Infect.* **149**, e240 (2021).
9. Castillo, O. et al. Interval type-3 fuzzy aggregators for ensembles of neural networks in COVID-19 time series prediction. *Eng. Appl. Artif. Intell.* **114**, 105110 (2022).
10. Qu, Z. et al. An innovative ensemble model based on multiple neural networks and a novel heuristic optimization algorithm for COVID-19 forecasting. *Expert Syst. Appl.* **212**, 118746 (2023).
11. Mohan, S. et al. An approach to forecast impact of Covid-19 using supervised machine learning model. *Softw.: Pract. Exp.* **52**(4), 824–840 (2022).
12. Dong, Y. et al. Evaluating the ability of the NLHA2 and artificial neural network models to predict COVID-19 severity, and comparing them with the four existing scoring systems. *Microb. Pathog.* **171**, 105735 (2022).
13. Shafiq, A. et al. Comparative study of artificial neural network versus parametric method in COVID-19 data analysis. *Results Phys.* **38**, 105613 (2022).
14. Nasiri, H. & Hasani, S. Automated detection of COVID-19 cases from chest X-ray images using deep neural network and XGBoost. *Radiography* **28**(3), 732–738 (2022).
15. Liu, T., Siegel, E. & Shen, D. Deep learning and medical image analysis for COVID-19 diagnosis and prediction. *Annu. Rev. Biomed. Eng.* **24**, 179–201 (2022).
16. Prezioso, E. et al. Predictive medicine for salivary gland tumours identification through deep learning. *IEEE J. Biomed. Health Inform.* **26**(10), 4869–4879 (2021).
17. Djavanshir, G. R., Chen, X. & Yang, W. A review of artificial intelligence's neural networks (deep learning) applications in medical diagnosis and prediction. *IT Prof.* **23**(3), 58–62 (2021).
18. Alakus, T. B., Turkoglu, I. Detection of pre-epileptic seizure by using wavelet packet decomposition and artificial neural networks. In *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*. IEEE, 511–515 (2017).
19. Memarian, N., Kim, S., Dewar, S., Engel, J. & Staba, R. J. Multimodal data and machine learning for surgery outcome prediction in complicated cases of mesial temporal lobe epilepsy. *Comput Biol Med.* **64**(1), 67–78 (2015).
20. Badkundri R, Valbuena V, Pinnamareddy S, et al. Forecasting the 2017–2018 Yemen cholera outbreak with machine learning. arXiv preprint arXiv. 1902.06739 (2019).
21. Abbasi, J. Better cholera counts through machine learning models. *JAMA* **321**(14), 1343–1343 (2019).
22. Asher, J. Forecasting Ebola with a regression transmission model. *Epidemics* **22**, 50–55 (2018).
23. Rachah, A. & Torres, D. F. M. Predicting and controlling the Ebola infection. *Math. Methods Appl. Sci.* **40**(17), 6155–6164 (2017).
24. Colubri, A. et al. Transforming clinical data into actionable prognosis models: machine-learning framework and field-deployable app to predict outcome of Ebola patients. *PLoS Negl. Trop. Dis.* **10**(3), e0004549 (2016).
25. Gong, J., Wu, Y. P. & Li, L. Parameters estimation in Ebola virus transmission dynamics model based on machine learning. *Physica A: Stat. Mech. Appl.* **536**, 122604 (2019).
26. Deng Q. Dynamics and prediction of the COVID-19 epidemics in the US: a compartmental model with deep learning enhancement. Available at SSRN 3622360. 2020.
27. Valsecchi, C. et al. Parsimonious optimization of multitask neural network hyperparameters. *Molecules* **26**(23), 7254 (2021).
28. Shen, M. et al. Nonlinear hyperparameter optimization of a neural network in image processing for micromachines. *Micromachines.* **12**(12), 1504 (2021).
29. Brodzicki, A., Piekarski, M. & Jaworek-Korjakowska, J. The whale optimization algorithm approach for deep neural networks. *Sensors.* **21**(23), 8003 (2021).
30. Aljarah, I., Farris, H. & Mirjalili, S. Optimizing connection weights in neural networks using the whale optimization algorithm. *Soft. Comput.* **22**, 1–15 (2018).
31. Alakus, T. B. & Turkoglu, I. Comparison of deep learning approaches to predict COVID-19 infection. *Chaos Solitons Fract.* **140**, 110120 (2020).
32. Mirjalili, S. & Lewis, A. The whale optimization algorithm. *Adv. Eng. Softw.* **95**, 51–67 (2016).
33. Pandey, A. C. & Tikkiwal, V. A. Stance detection using improved whale optimization algorithm. *Complex Intell. Syst.* **7**, 1649–1672 (2021).
34. Nadimi-Shahraki, M. H., Taghian, S. & Mirjalili, S. An improved grey wolf optimizer for solving engineering problems. *Expert Syst. Appl.* **166**, 113917 (2021).

35. Malathi S, Raj Y, Arockia K, *et al.* Prediction of cardiovascular disease using deep learning algorithms to prevent COVID-19. *J. Exp. Theor. Artif. Intell.* 1–15 (2021).
36. Kwon, B. C. *et al.* Retainvis: Visual analytics with interpretable and interactive recurrent neural networks on electronic medical records. *IEEE Trans. Visual Comput. Graphics* 25(1), 299–309 (2018).
37. Xu, Y. *et al.* Deep learning predicts lung cancer treatment response from serial medical imaging longitudinal deep learning to track treatment response. *Clin. Cancer Res.* 25(11), 3266–3275 (2019).
38. Yongyi, Z. H. A. O., Fei, W. A. N. G. & Ying, S. H. E. N. Prediction of COVID-19 epidemic trend sequence analysis based on long short-term memory network. *J. Shenyang Normal Univ. (Natural Science Edition)*. 38(06), 525–531 (2020).
39. Wu, Q. *et al.* Prediction of epidemic situation in COVID-19 based on time series neural network. *J. Anhui Univ. Technol. (Natural Science)* 38(02), 188–194 (2021).
40. Cheng, N., Ding, C., Gao, W. & Liu, J. Analysis of COVID-19 outbreak based on window-time-LSTM model. *Chin. J. Dis. Control Prevent.* 25(05), 577–582 (2021).
41. Kim, I., Rajaraman, S. & Antani, S. Visual interpretation of convolutional neural network predictions in classifying medical image modalities. *Diagnostics*. 9(2), 38 (2019).
42. Iglesias, L. L. *et al.* A primer on deep learning and convolutional neural networks for clinicians. *Insights Imaging* 12(1), 1–11 (2021).
43. Schwab, P., Schütte, A. D., Dietz, B., *et al.* PredCOVID-19: a systematic study of clinical predictive models for coronavirus disease 2019. arXiv preprint arXiv. 2005.08302. 76 (2020).
44. Hajian-Tilaki, K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J. Internal Med.* 4(2), 627 (2013).
45. Kamarudin, A. N., Cox, T. & Kolamunnage-Dona, R. Time-dependent ROC curve analysis in medical research: Current methods and applications. *BMC Med. Res. Methodol.* 17(1), 1–19 (2017).
46. Zongxi, Q., Yongzhong, S. & Yutong, L. Predicting major infectious diseases based on grey wolf optimization and multi-machine learning: case study of COVID-19. *Data Anal. Knowl. Discov.* 6(8), 122–133 (2022).
47. Liu, D. *et al.* Optimizing deep neural networks to predict the effect of social distancing on COVID-19 spread. *Comput. Ind. Eng.* 166, 107970 (2022).
48. Singh, P., Chaudhury, S. & Panigrahi, B. K. Hybrid MPSO-CNN: Multi-level particle swarm optimized hyperparameters of convolutional neural network. *Swarm Evol. Comput.* 63, 100863 (2021).
49. Sun, F. *et al.* The prediction of new medical resources in China during COVID-19 epidemic period based on artificial neural network model optimized by genetic algorithm. *J. Phys.: Conf. Ser.* 1815(1), 012033 (2021).
50. Qian, L. I. *et al.* Modelling COVID-19 epidemic with time delay and analyzing the strategy of confirmed cases-driven contact tracing followed by quarantine. *Acta Math. Appl. Sin.* 43(2), 238–250 (2020).
51. Pasaribu, U. S. *et al.* Modelling COVID-19 growth cases of provinces in Java Island by modified spatial weight matrix GSTAR through railroad passenger's mobility. *Heliyon*. 7(2), e06025 (2021).

## Author contributions

X.S.: Conceptualization, Data curation, Formal analysis, Methodology, Writing—Original draft, Writing—Review and editing, Y.S.: Data curation, Formal analysis, Methodology, Writing—Review and editing, Supervision, H.L.: Formal analysis, Methodology, Software, Q.L.: Data curation, Formal analysis, Y.Z.: Formal analysis, Methodology, Writing—Review and editing, J.Z.: Writing—Review and editing, Supervision, \* C.W.: Data curation, Formal analysis, Methodology, Writing—Review and editing, Supervision, Y.C.: Writing—Review and editing. All authors have read and agreed to the published version of the manuscript.

## Funding

This research was funded by Jilin Province Science and Technology Development Plan Project of China (20230201091GX); Jilin Provincial Natural Science Foundation of China (20230101237JC); the Science and Technology Development Planning of Jilin Province (20210203006SF); the sixth batch of Jilin Province youth science and technology talent promotion project (QT202216).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to C.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023