An Evaluation Method of English Composition Automatic Grading Based on Genetic Optimization Algorithm and CNN Model

Li Wang
School of Foreign Studies, Henan University of Urban Construction, Pingdingshan 467000, China
University of the Cordilleras, Baguio City,2600, Philippines

***Abstract*—In response to the problems of traditional genetic algorithms in evaluating English compositions, the stability of automatic grading of English compositions has been further enhanced. This article evaluates the teaching effectiveness of automatic grading of English compositions using an optimization fusion algorithm combined with genetic optimization algorithm and CNN model. By analyzing genetic content and optimization algorithms, a corresponding fusion optimization model was obtained, and the automatic evaluation of English compositions was analyzed and predicted through experimental verification. The results indicate that the curves corresponding to different parameters exhibit typical segmentation features through the variation curves of individual numbers under different scale factors. And through quantitative description and analysis of the curve, it can be seen that the change in proportion factor has an absolute advantage in the impact of genetic algorithm on the number of children. As the number of samples increases, the performance of genetic optimization algorithms under the f function shows an upward trend. Research has shown that the writing content index has the greatest impact on English writing, while the corresponding grammar errors have the smallest impact on English writing. Finally, the accuracy of the optimized model was verified by comparing the model curve with experimental data. This study provides theoretical support for the use of genetic optimization algorithms and CNN models in English, and provides ideas for the use of optimization algorithms in other fields.**

Keywords—Genetic optimization algorithm; CNN model; English composition; Automatic scoring; Teaching effect

I. INTRODUCTION

Researchers have proposed various evaluation methods for the quality of English teaching in universities, such as grey relational analysis and fuzzy comprehensive evaluation [1]. However, these methods are suitable for linear models and are difficult to adapt to nonlinear teaching quality evaluation problems. They have subjective and random defects and cannot effectively achieve teaching quality evaluation. In terms of natural language processing models, researchers mainly use neural network-based models and convolutional neural network-based models. These models can convert text information into vectors and perform part of speech analysis, word form restoration, and other processing to ultimately obtain a feature vector related to text information [2]. In terms of scoring algorithms, researchers mainly used machine learning based algorithms and deep learning based algorithms. Among them, machine learning based algorithms mainly improve the accuracy of prediction by training models [3]. The algorithm based on deep learning mainly falsely optimizes the structure and weight of the neural network to improve the accuracy of prediction [4]. The evaluation of teaching quality for university teachers is an important way to improve teaching management level and teaching ability of teachers. By utilizing teaching quality evaluation, students can provide feedback on the teacher's teaching situation, and teachers can reflect on the teaching effectiveness. Schools can effectively implement teaching management improvements and provide targeted training for teachers. English teaching is an important part of higher education, and the evaluation process of English teaching quality is relatively complex. Therefore, constructing an objective and scientific evaluation model for English teaching quality is a hot research direction [5]. Given the problems existing in chemical reactions, genetic optimization algorithm was used to improve the accuracy of chemical reduction reactions [6]. Firstly, the chemical reaction data were imported into the optimization model, and then the characteristic parameters of the test data were derived through further analysis of the model. Finally, the optimized model was used to modify the derived information. To verify the correctness of the model, relevant experiments were applied to verify and analyze the results of chemical reactions. In order to further improve the quality of tea, the traditional genetic algorithm can be optimized to reflect and describe the relevant properties of tea [7]. Different evaluation indexes can be obtained by extracting the characteristic parameters of tea leaves. Through the arrangement and analysis of the evaluation indexes, the optimized index data can be obtained, and the quality of tea can be judged through the analysis of the data. Finally, a large number of data were used to evaluate the correctness of the model. The convolutional neural network also has applications in various domains: Aiming at the issues of rock identification, the convolutional neural network can be modified to obtain the optimized convolutional neural network. The optimization model can better identify and analyze the rock [8]. The corresponding influencing factors can be found by extracting the rock's data features. Thus, the variation curves of rock characteristics under different influencing factors were obtained. Experiments were carried out to verify the model. The results show that the model can describe the characteristic parameters of rock well. Convolutional neural networks can also play a greater role in medical monitoring [9]. Specific medical parameters can be obtained through scanning and extracting samples, different etiologies can be studied through description and analysis, and accurate results can be finally obtained.

Lightweight encryption (LWC) is an emerging technology used to develop encryption algorithms or protocols for implementation in restricted environments, such as WSNs, RFID tags, smart medical devices, and many other embedded systems. It is expected that LWC will play an important role in ensuring the Internet of Things and universal computing. The term 'lightweight' can be considered from two perspectives, namely hardware and software. However, portability in hardware does not necessarily mean portability in software, and vice versa. The above studies were mainly from medicine, architecture, and so on, but applying genetic algorithms in English composition was relatively rare. To further improve the application of the genetic algorithm and CNN model in English composition evaluation, based on the genetic theory, the fusion optimization model of the genetic algorithm and CNN theory was adopted to study the automatic scoring of English composition. The evaluation rules of scoring effects under different factors were obtained by analysing English composition's characteristics and related indicators. Eventually, the accuracy of the model was confirmed by associated experiments. This optimization model can improve the research ideas for applying genetic optimization algorithms and the CNN model in other fields.

This article optimizes the problems that exist in traditional evaluation of English compositions. The research has the following innovations:

1) The method proposed in this article is suitable for linear models and for teaching quality evaluation problems with nonlinearity. It solves the defects of subjectivity and randomness, and can effectively achieve teaching quality evaluation.

2) The corresponding fusion optimization model was validated through experiments to analyze and predict the automatic evaluation of English compositions.

As the number of samples increases, the performance of genetic optimization algorithms under the f function shows an upward trend. The curve corresponding to the g function is stable.

Section I analyzes the relevant background of automatic grading and evaluation of English compositions using genetic optimization algorithms and CNN models. Section II analyzes the relevant theories of genetic optimization algorithms. Section III further analyzes and understands the computational process and structural unit of convolutional neural networks. By organizing and analyzing relevant data, a convolutional neural network structure diagram was drawn. Section IV elaborates on the application of genetic optimization algorithms and CNN models in English writing, and analyzes the basic content of the evaluation method for the teaching effectiveness of automatic grading of English writing. The model curve in Section V indicates that the optimized model can effectively illustrate and characterize the trend of changes in English composition indicators. Therefore, in order to study the impact of automatic grading of English compositions on teaching, an optimization model can be used.

II. Related Theory of Genetic Optimization Algorithm

A biological individual is considered a response to the optimization algorithm in a genetic algorithm, a random search algorithm [10]. The population, which consists of many individuals, is the algorithm's solution set. Next, the weaker individuals are continuously eliminated through genetic processes like selective crossover and mutation. The genes with the favorable mutation are passed on to the next generation through probabilistic choice. The output individual is the ideal one that endures after screening when the termination condition is reached after continuous iteration. Both discrete and continuous optimization problems can be solved using the genetic algorithm, which is simple to use. Its multi-direction global optimization performance has good theoretical value for addressing challenging optimization problems today.

The existing cryptography relies on mathematical algorithms and key size, which can take decades to hundreds of years to crack. In theory, quantum computing can break existing encryption faster, possibly within days or even minutes. PQC is a new encryption method that resists quantum computing attacks by using very difficult mathematical problems that quantum computers cannot solve in a reasonable amount of time.

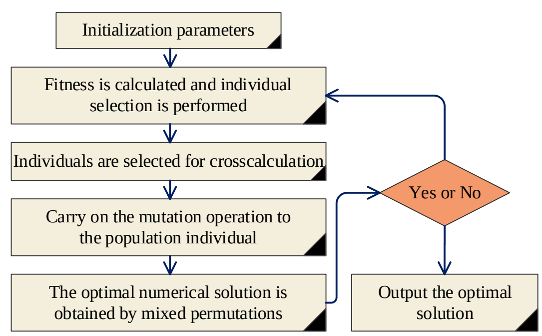

Fig. 1. Genetic algorithm solution flow chart.

Side channel attack is a cryptographic analysis method that uses information other than the encryption algorithm itself to infer certain information in the encryption algorithm, which may lead to the leakage of encrypted data. Power analysis uses the changes in power consumption during the encryption process to derive the state of the encryption algorithm and then crack the key. Time analysis uses the execution time differences of different encryption operations to infer the internal state of the encryption algorithm. Lightweight cryptography is a cryptographic method targeting resource constrained environments such as IoT devices, mobile devices, and embedded systems. Due to the limited computing power, storage resources, and energy supply in these environments, lightweight cryptography requires less computational and storage resources, while ensuring the security and integrity of encrypted data. To analyze the solving process of the genetic optimization algorithm, it summarized and analyzed different genetic algorithms and thus obtained the flow chart of genetic optimization calculation under other algorithm theories as shown in Fig. 1. Considering the calculation process in the figure, initialization analysis should be carried out on relevant data first, making the study results more accurate. Initialization parameters include population size, iteration times, etc. The features of parameters can be extracted through initialization, and then the extracted feature parameters can be imported into the next calculation. In the next step, individuals can be selected through fitness calculation to obtain optimized genetic optimization data. Then, it is imported into the genetic optimization algorithm for cross-calculation, and the characteristics of the population are optimized and modified through cross-calculation. Then, the relevant data of population quantity are imported into the variation feature module for mutation operation, and then the data obtained from mutation operation are mixed and arranged, and the optimal number solution is obtained through calculation. Finally, judge whether the termination condition is met. If not, further iterative analysis is required. If so, the data will be output.

A. Genetic Algorithms

Genetic algorithm is an optimization algorithm that simulates the process of biological evolution. It seeks the optimal solution by simulating processes such as gene combination, crossover, and mutation. In networks that require multi-agent management, genetic algorithms can be used to optimize transmission quality. For example, by simulating the interaction and competition process between agents in a network, genetic algorithms can automatically find an optimal network transmission strategy [11]. The process of using a genetic algorithm mainly involves two aspects such as cross-real number recombination and mutation [12], [13]. (1) Crossover and recombination of real numbers: the commonly used methods mainly include discrete and intermediate recombination in the recombining real numbers.

1) Discrete recombination: The operation of discrete recombination is relatively simple; that is, the genes on the above individuals before optimization are equally probabilistic and randomly inherited into the optimized individuals to realize individual renewal.

2) Intermediate recombination: The intermediate recombination completes the individual update process by calculating the expression as follows:

$$\mathrm{X}=\mathrm{X}_1 + a(X_2 - X_1) \quad (1)$$

X represents the child produced, X_1 and X_2 are the two selected parents, and a is the scaling factor.

Different scale factors influence the number of individuals in the genetic algorithm. To evaluate the influence of scale factors on the number of individuals, it draws the change curves of the number of individuals under different scale factors. Fig. 2 reveals that the number of individuals of the scale factor corresponding to additional parameter a shows different variation trends, as shown below: When a=-1, the corresponding curve shows typical two-stage variation features. In the first stage, the curve shows a gradual decline as the number of iterative steps increases, and the curve slope indicates that the curve at this stage is linear. However, when the number of iterations exceeds 60, the corresponding curve manifests an approximate constant trend of change, indicating that the number of sub-individuals decreases gradually with the increase of the number of iterations. A two-stage variation trend is still visible in the corresponding curve when parameter a equals to 1. With more iterations in the first stage, the curve gradually gets better. As opposed to this, the slope of the corresponding curve is greater than that of the parameter a = -1, demonstrating that the influence of the parameter a = 1 on the number of sub-individuals is greater than that of the parameter a = 1. The corresponding curve shows a slight or gradual decline when the number of iterations exceeds 60. The overall curve thus demonstrates that when parameter a = 1, the influence of the number of corresponding sub-individuals still affects the change in volatility as a whole. When a=0, the associated curve exhibits a trivial change trend, and when its scaling factor is greater than 60, the related curve shows a slight growth trend. The number of sub-individuals under the three parameters above demonstrates that the influence of neutron individual numbers has a clear advantage in changing the scale factor. Therefore, when calculate the scale factor of the genetic optimization algorithm, it need to consider the specific variation of parameters.

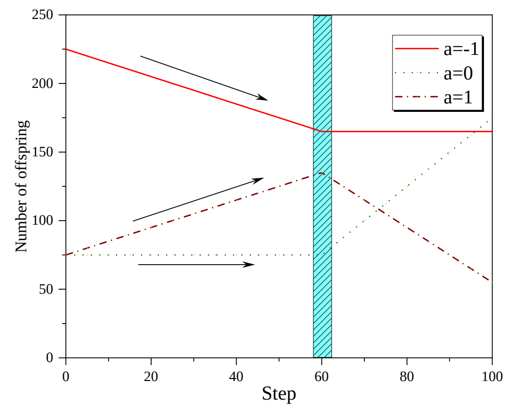

Fig. 2. Scale factor variation diagram.

3) Variation: Variation refers to the change of some genes in the calculation process of the genetic optimization algorithm so that the characteristics of the number of offspring are different from those of other individuals. The variation of the genetic optimization algorithm can be divided into real number variation and binary variation according to separate research contents and calculation methods.

a) Real number variation: The following formula is generally used for real number variation.

$$X_3=\mathrm{X} \pm 0.5L\Delta \quad (2)$$

$$\Delta = \sum_{\mathrm{i}=0}^{m} \frac{a(i)}{2^i} \quad (3)$$

X_3 is the individual after mutation, and L is the value range of the variable. △ is the coefficient of variation, m is an artificially set integer, and a (i) is the probability function.

b) Binary Variation: Binary variation is relatively simple. For the binary code of characters, several positions are randomly selected on the individual gene sequence first, and then the gene values on these loci are reversed.

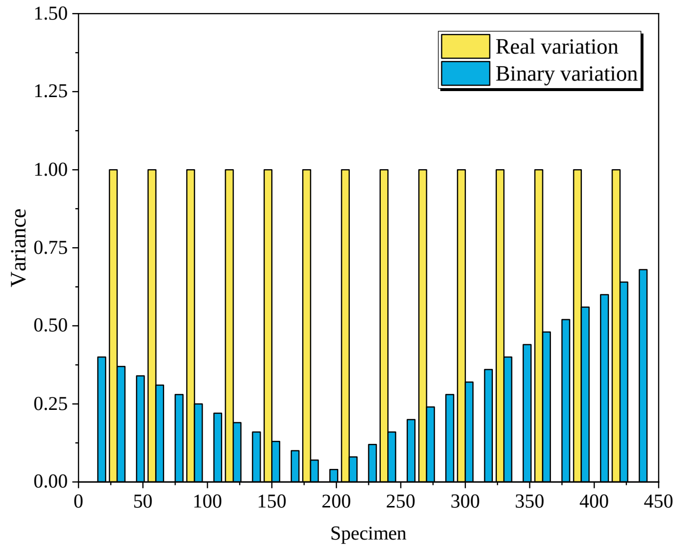

Fig. 3. Diagram of changes of two variation indicators.

Through variation calculation, it can see that different calculation methods have other influences on the number of variations. To study the effect of real number variation and binary variation on the number of variations, the change curves under two variation indicators were drawn, as shown in Fig. 3. As the figure displays, with the gradual increase in the number of samples, the two different variation modes show typical variation trends. Firstly, it can be seen from the real number variation method that the real number variation shows a distinct piecewise change as the number of iterations increases. When the number of samples is odd, the corresponding sample variation is 1. When the corresponding variation sample reaches an even number, the corresponding sample variation is zero. Therefore, the real number variation shows the typical characteristics of the 01 variation. As can be seen from the curve corresponding to the binary variation method, when the variation gradually increases, the corresponding number of variations shows a gradual decline, and the slope of the curve corresponding to the downward trend is approximately constant. When the number of samples exceeds 200, the corresponding curve gradually increases. Through the above analysis, it can see that binary variation shows different trends with different sample numbers. Therefore, relevant studies show that to improve the accuracy of the genetic optimization algorithm, it is necessary to use a wide range of samples for iteration and analysis.

B. *Genetic Algorithms Update Content*

It can be seen from the above research that the genetic algorithm conducts targeted research on input data through real number variation and binary variation and finally outputs neurons through an iterative updating calculation method [14], [15]. Therefore, the individual update method can be adopted to optimize the genetic algorithm, and the corresponding optimization and update steps are as follows:

1) Determine the neural network of the genetic algorithm: To solve the neuron j corresponding to the initialisation data of the optimization problem and the corresponding weight vector W_j, the connected structural network of the genetic algorithm should be adopted [16], [17].

2) Select individual samples from the individuals: Produced by the genetic algorithm as input data into the genetic algorithm structural network. The corresponding calculation formula is as follows:

$$\mathrm{dis}(W_j, X_t) = \sqrt{\sum_{k=0}^{k=n}(w_j^k - x_j^k)^2} \tag{4}$$

where is the gene value at the kth position on the j-th neuron, x_j^k represents the gene value at the kth position on the i-th individual, and X_t is the sample vector.

3) Solve the best matching unit: determine the relationship and distance between the genetic algorithm data in the above way, and find the neuron data with the closest distance.

4) Calculate the neighborhood radius of the neuron:

$$\begin{cases} \sigma=\sigma_0\exp(-t/\lambda) \\ \lambda = N_0/\,lg(\sigma_0) \end{cases} \tag{5}$$

Where is the initial field radius; λ is the time constant?

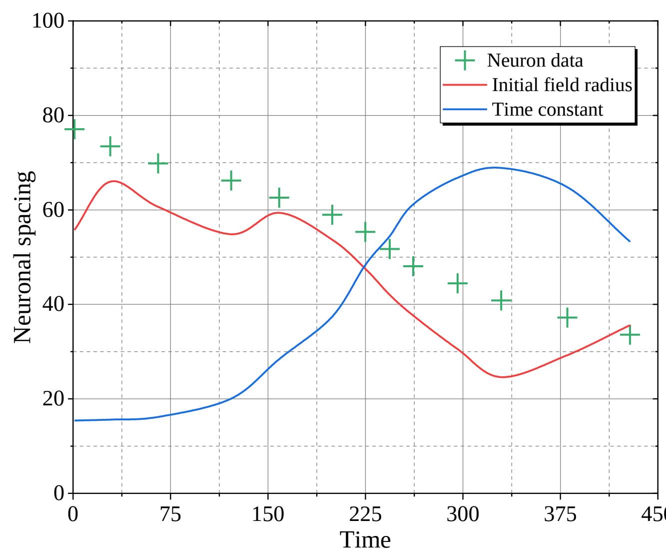

Fig. 4. Effect of the coefficient on neuronal spacing.

According to the above formula and analysis, it can be seen that the specific parameters of neurons greatly influence the spacing of neurons, and the different parameters will lead to differences in the calculation methods and theories of the spacing of neurons. To study the influences of two factors on neuron spacing, it drew the variation curves of neuron spacing of the genetic optimization algorithm under two parameters, as shown in Fig. 4. It can be seen from the curves that the variation trend of neurons under two different parameters presents typical nonlinear changes. In contrast, the corresponding neuron curve shows a gradual downward trend. Moreover, it can be seen from the slope that the linear characteristics of neuron data are obvious. Specifically, first of all, it can see through the initial field radius curve rise gradually. The linear decline, when it reaches the minimum spacing of neurons and shows the increasing trend, gradually declines, then hits the stage when it rises again, showing the typical volatility change trend. As shown in the influence curve of the time constant, the corresponding curve first exhibits a stable change trend as time increases gradually, and then the corresponding curve gradually increases. And when it reaches its highest point, the corresponding curve slowly decelerates. The slope of the corresponding curve also exhibits a gradual increase trend. The overall data feature of the curve with time parameter has the comprehensive effect of linear and nonlinear. Therefore, it must comprehensively consider the impact of the initial domain radius and time constant on neuron spacing.

5) Solve the neuron update formula: Through the above analysis and solution, the influence of parameters on the spacing between different neuron data can be obtained, and the corresponding update formula is as follows:

$$w_j^k(\mathrm{t+1})=\mathrm{w}_j^k(t)+\mathrm{L}(t)\theta(t)(x_j^k(t)\text{-}\mathrm{w}_j^k(t)) \quad (6)$$

where is the gene value at the position after the update, is the learning speed factor, and is the best Euclidean distance?

6) Determine whether the termination conditions are met: The neuron data are compared by setting the termination conditions of the corresponding predictive optimization algorithm.

7) Update individual data and location information. The corresponding update formula is as follows:

$$\mathrm{X}=\mathrm{w}^* + \sum_{\mathrm{k=0}}^{\mathrm{m\text{-}1}} a_j^k U_j^k + \mathrm{N}(0,\sigma\mathrm{I}) \quad (7)$$

Where represents the best matching unit with the individual; a_j^k is the Euclidean distance between adjacent spirit longitude elements; U_j^k is the unit vector of adjacent neurons; m is the number of targets; U_j^k is a normally distributed noise vector; I am the identity matrix.

For analyzing the impact of various parameters in the update formula on the updated data of the genetic optimization algorithm, relevant data were obtained through iterative calculation. The influence curves of four different parameters on the genetic algorithm were drawn, as shown in Fig. 5. Different parameters exhibit typical variation characteristics, as can be seen in the figure. As can be seen, the updated data for the corresponding parameter a shows a trend of slow increase with the gradual increase of iteration times and time. However, a slight nonlinear variation can be seen on the corresponding fitting curve. As shown by the parameter U, the curve initially declines gradually. When there are more than four iterations, the corresponding appropriate curve initially shows a nonlinear change before fluctuating and exhibiting a sharp downward trend. It can be seen from parameter m that the overall curve remains at about 70, indicating that parameter m has the least influence on genetic algorithm update. Still, it also shows that the parameter has a relatively clear specific value of genetic algorithm update and good corresponding stability. It can be seen from parameter I that the curve shows typical linear characteristics, the corresponding quasi-sum curve has obvious smoothness, and the corresponding data presents a trend of a gentle rise. Through the above analysis, various parameters have different influences on the genetic algorithm, and it is necessary to select specific parameters according to the actual situation to analyze the genetic optimization algorithm to acquire precise outcomes.

C. Test Function Selection of Genetic Optimization Algorithm

Since the complexity of optimization problems affects the computational efficiency of swarm intelligence algorithms [18], [19], it is necessary to select four commonly used test functions, and their corresponding forms are shown as follows:

1) Test function f(x): Test function f(x) is an optimised test function with a simple structure consisting of two variables and a constraint condition.

$$f(x)=\mathrm{x}_1 + (x_2\text{-}1)^2 \quad (8)$$

2) Test function g(x): This function is a test function of multidimensional single constraint conditions.

$$g(x) = (\sqrt{n})^n \prod_{i=1}^{n} x_i \quad (9)$$

3) Test function h(x): Test function H (x) is a low-dimensional multi-constraint single-objective optimization problem.

$$h(x) = \frac{\sin^3(2\pi\mathrm{x}_1)\sin^2(2\pi\mathrm{x}_2)}{x_1^3(x_1+\mathrm{x}_2)} \quad (10)$$

4) Test function k(x): This function is a multi-dimensional and multi-constraint complex nonlinear optimization problem, which is representative of verifying the algorithm's feasibility.

$$k(x) = (x_1 - 10)^3 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 10)^2 \quad (11)$$

It can be seen from the above analysis that genetic optimization algorithms contain different test functions, and the other test functions lead to different solving processes and methods for nonlinear problems of genetic algorithms. To analyze the influence of the four tests performs on the performance of the genetic algorithm, the variation curves of the genetic optimization algorithm under different samples were obtained by summarizing the data, as revealed in Fig. 6. Regarding the change curve in the figure, it can be seen that the performance of the corresponding genetic optimization algorithm shows a gradually increasing trend with the gradual growth of the number of samples of the f function. However, when the number of samples is about 65, the corresponding genetic optimization performance data shows a sudden decline because when the number of samples is 65, the corresponding calculation method will further propose the characteristics of this parameter. As a result, the corresponding genetic optimization performance index has declined. From the change rule of g performance, it can be seen that the corresponding bar chart has a constant adjusted trend. And the slope is the same up and down. It shows that the genetic optimization performance of the g function is stable. It can be seen from the change of the h function that the corresponding genetic optimization performance index decreases gradually from the maximum to the minimum, which indicates that the data of the genetic optimization performance corresponding to this function has a large fluctuation, showing a trend of gradual decline. Finally, it can be seen from the change of the k function that the corresponding data index shows an incremental improvement. When the corresponding data reaches the maximum, it indicates that the data of the genetic optimization algorithm of the corresponding k function shows a gradual increase trend of change.

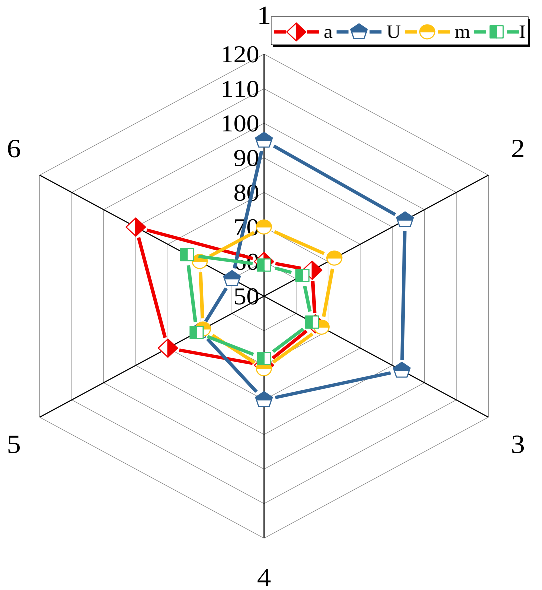

Fig. 5. Influence of different parameters on genetic algorithm update.

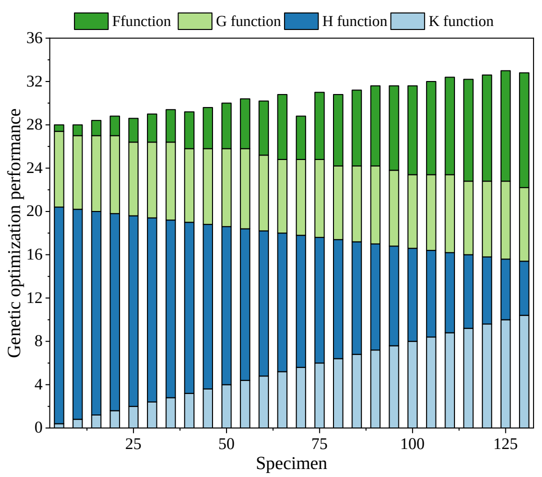

Fig. 6. Test function variation diagram.

III. CNN Model

A convolutional neural network is a typical structure of artificial neural networks, which has the characteristics of local connection, weight sharing and pooled sampling and has gradually become a representative network for feature extraction in deep learning networks [20], [21].

To further analyze and understand the convolutional neural network's computation process and the structural unit, the convolutional neural network, the structure diagram of the convolutional neural network was drawn by sorting out and analyzing relevant data, as shown in Fig. 7. It can be seen from relevant studies that convolutional neural networks can be divided into five modules according to different research contents and methods; corresponding contents mainly include the convolution layer, pooling layer and full connection layer. In the corresponding computing module of the convolutional neural network, the model data is first imported into the specific computing module. Then the data features are analyzed and processed by further data feature extraction module. The optimised results are output through data optimization and modification analysis, and experimental data further verify the optimized data. Thus, the optimized convolutional neural network calculation results are obtained. The main calculation steps of the convolutional neural network are as follows:

1) Similarity measurement and analysis: In the similarity measurement process of the convolutional neural network, the correlation distance formula is widely used in similarity measurement, and the corresponding procedure is shown as follows:

$$d(\mathrm{x,y}) = (\textstyle\sum_{\mathrm{i=1}}^{n} \left|x_i\text{-}\mathrm{y}_i\right|^{p})^{1/\mathrm{p}} \quad (12)$$

where x and y respectively represent the vector representation of two image samples, n represents the length of the vector representation, and p is a constant.

The similarity measurement method of the i-th element can be measured as follows:

$$z_i = \frac{1}{h\times w}\textstyle\sum_{\mathrm{i=1}}^{h}\sum_{\mathrm{j=1}}^{w} F(\mathrm{i,j}) \quad (13)$$

where h represents the height of the feature graph, w represents the width of the feature graph, and F represents the channel of the feature graph.

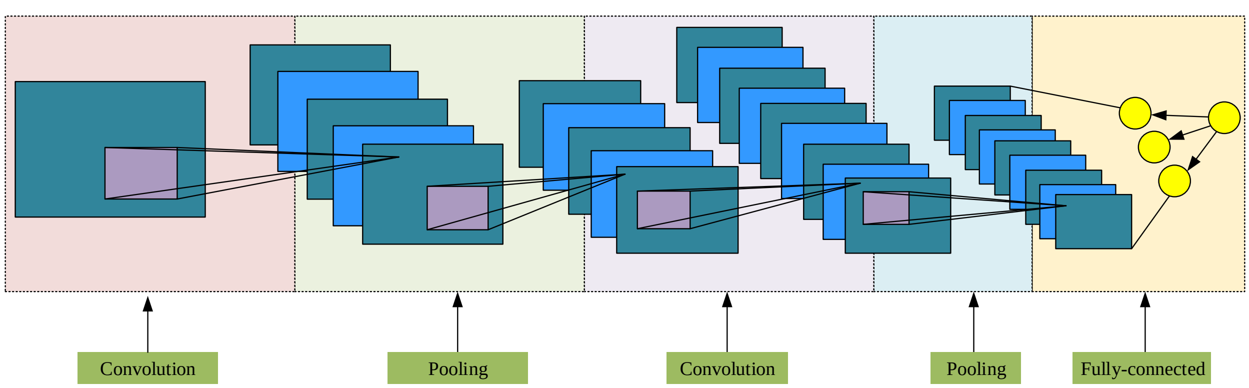

Fig. 7. Structure diagram of convolutional neural network.

2) Sample generation: When the directional pin-mixing augmented method is used to analyze the genetic optimization algorithm, the corresponding implementation process is shown as follows:

$$\begin{cases} \mathrm{I}=(1-\mathrm{M}_\lambda) \odot I_m + \mathrm{M}_\lambda \odot I_m \\ \mathrm{y}=\lambda \mathrm{y}_m + (1-\lambda \times y) \end{cases} \tag{14}$$

where I represent sequence mixing and y represents sequence category label. Describes the specific gravity of data, I'm representing the sequence of training samples, and y_m represents the sequence of category tags. It is the dot product operation.

3) Training sample evaluation: In data recognition of convolutional neural network, classification accuracy R is generally used as the evaluation result, and the corresponding formula is shown as follows:

$$\mathrm{R}=\frac{I_r}{I} \tag{15}$$

I_r represents the data quantity of classification, and I represent the data quantity of total evaluation.

To analyze the influence of convolutional neural network with accuracy index, the accuracy change curves under different iterations and accuracy index were drawn through experimental analysis, as shown in Fig. 8. The research shows that the accuracy index and rate are typical two-stage structures. In the first stage, the corresponding accurate index I_r shows a trend of gradual decline. And its slope remains constant, indicating that the data of the precise index shows a linear change characteristic, while the same accurate total I show a gradually increasing trend with the increase of the number of iterations. This is because the accuracy index increases with the rise of characteristic parameters corresponding to the number of iterations, leading to an increase in calculation results. It can see an obvious trend of gradual decline through the curve of the corresponding accuracy R. This indicates that when the number of iterations is from 0 to 100, the corresponding accuracy rate decreases gradually with the increasing number of iterations. In the second stage of the curve, with the further increase in the number of iterations, the corresponding accurate index shows the same linear change. It belongs to the evolution of linear increase. The total number of corresponding indicators declined rapidly with the rise in iteration times, and the corresponding downward trend showed an obvious linear decline. As can be seen from the accuracy curve, with the increase in the number of iterations, the related accuracy shows a gradually increasing trend. This indicates that when the number of iterations exceeds 100, the accuracy of the connected convolutional neural network gradually improves. Therefore, in the actual use and calculation process, it needs to iterate and calculate the genetic optimization algorithm and convolutional neural network many times so as to obtain more accurate results.

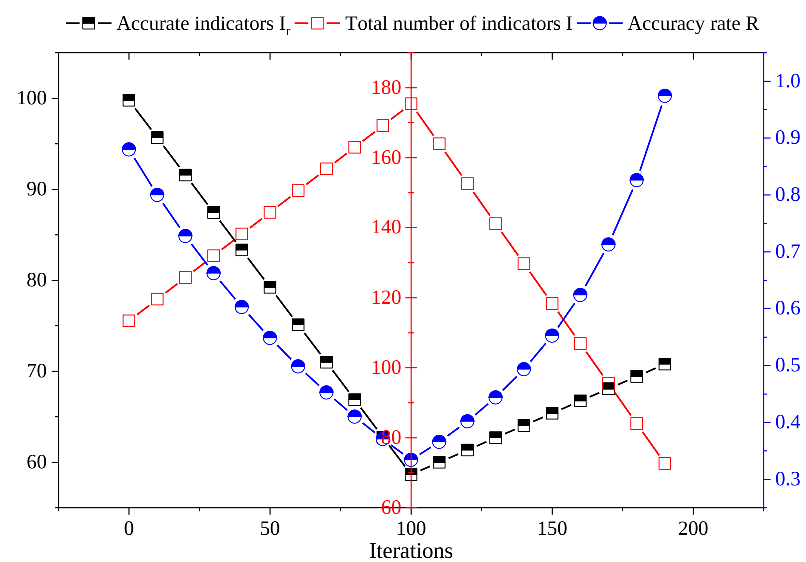

Fig. 8. Change in accuracy chart.

IV. Application of Genetic Optimization Algorithm and CNN Model in English Composition

A. *The basic Content of English Composition Automatic Grading Teaching Effect Evaluation Method*

English composition is very important for English teaching and assessment. Still, there are a series of problems in the practical application of English composition evaluation, mainly including: (1) English composition scoring efficiency is low: In the useful application process, the efficiency of English composition is low because of the artificial evaluation. (2) The standard of English composition scoring is not unified: There is no unified standard for English composition evaluation, which is greatly influenced by human factors in the actual evaluation process, resulting in the corresponding English composition evaluation results are not unified. (3) Single evaluation method for English compositions: English compositions are analyzed by single evaluation index in the evaluation process, resulting in the corresponding evaluation results are too single.

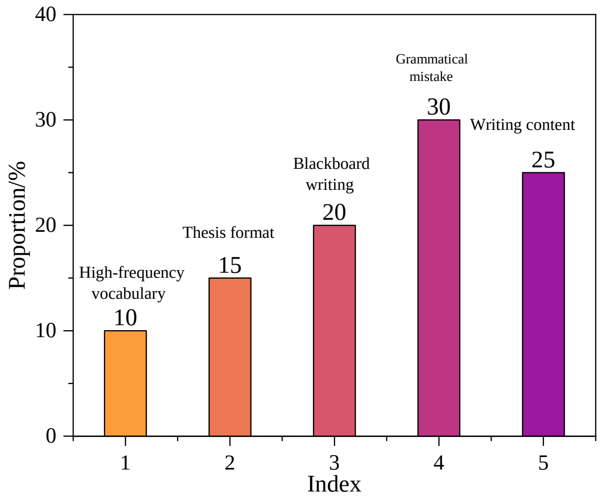

Fig. 9. English composition automatic scoring index distribution map.

To further analyze the influence of different indicators of automatic English composition scoring on English teaching effect, five various indicators are obtained through statistical analysis, including 1) high-frequency vocabulary, 2) composition format, 3) blackboard writing, 4) grammatical errors and 5) writing content, to analyze further the problems existing in self-scoring of English composition, the distribution chart of the comprehensive scoring index of English composition was obtained by summarizing and analyzing different data, as shown in Fig. 9. Regarding the results in the figure, the proportion of grammatical errors is the highest (30%), the content of writing is 25%, the handwriting on the blackboard is 20%, the composition format is 15%, and the proportion of high-frequency words is the lowest (10%). The original CNN neural network has a high accuracy in evaluating the quality of English teaching, both exceeding 81%. However, the evaluation accuracy of genetic algorithm optimized CNN neural network is better than that of the original CNN neural network The statistical results show that in the test samples, the genetic algorithm optimized CNN neural network model has an evaluation accuracy of more than 90%,. And some of the group evaluations have accuracy greater than 93%, indicating that the model has high approximation accuracy.

To further test the performance of the CNN neural network based on genetic algorithm optimization for English teaching quality evaluation constructed in this article, it is compared with the GA-BPNN evaluation model and the support vector machine (SVM) based evaluation model for performance testing. The simulation data remains unchanged.The experimental results show that the genetic algorithm optimized CNN neural network model constructed in this paper has high evaluation accuracy and operational efficiency This is because there is overfitting in the BP neural network in the GA-BPNN evaluation model, which affects the accuracy of the evaluation. In the SVM evaluation model, there are too many evaluation indicators, which interfere with each other, affecting the accuracy of the evaluation and increasing the computational cost.

B. Results

The above analysis and research show many problems applying the teaching effect of automatic scoring English composition. According to the related theory of genetic algorithm, a new optimization model is adopted to evaluate and analyze the teaching effect of automatic scoring of English composition based on the genetic optimization algorithm and CNN algorithm to further analyse the indicators related to automatic scoring of English composition. Thus, the flow chart of automatic scoring of English composition under the combined action of the genetic optimization algorithm and CNN model is obtained, as shown in Fig. 10. The specific calculation process is as follows: First of all, the basic calculation parameters of the genetic optimization algorithm and CNN fusion model are also set, and then relevant imported data are extracted specifically to obtain different types of data features. Then the automatic scoring scheme of English composition is analyzed and determined according to the various data characteristics. Then the relevant English data is imported into other genetic optimization algorithm calculation modules on the basis of determining the scheme. It may be divided into three diverse modules, and the contents of each type of module are basically the same, including data import, extraction of corresponding parameters, update of English indicators, discrimination of corresponding standards and output of final results. The data calculated by different modules are imported into the final local termination condition, and the data can be decided whether to continue the iteration or directly output the results through the discrimination of the termination condition.

By adopting the fusion optimization model of the genetic optimization algorithm and CNN model calculation method, the evaluation analysis diagram of the automatic scoring teaching effect of English composition under different indexes was finally obtained, as shown in Fig. 11. It can be seen from the variation trend of other indicators in the figure that the increase in time greatly influences the proportion of indicators. The specific changes are as follows: With the gradual increase of high-frequency words over time, the proportion of the corresponding English composition index displays a slight decline, and the nonlinear characteristics of this trend are obvious. This indicates that the influence of high-frequency words on the evaluation effect of English compositions is relatively simple, and this index cannot be used only to evaluate and analyze English compositions. With the increase of time, the corresponding composition format shows a fluctuating change; that is, it increases slowly at first and then decreases gradually. The writing index on the blackboard shows a relatively stable trend at different times, which indicates that the writing index is less affected by time, indicating that this index is an inherent attribute of English composition, which can be used to evaluate the teaching effect of English composition scoring under the genetic optimization algorithm and CNN model. The variation trend of grammatical errors is basically the same as that of blackboard writing. Still, the corresponding data is less than that of blackboard writing, indicating that grammatical errors' influence on English composition is less than that of blackboard writing. Finally, through the specific content index of the paper, it can be seen that with time, the proportion of English composition as a whole shows a gradually increasing trend. And the corresponding data are higher than the related data of other indicators in the same period. This indicates that the writing content index influences English writing most among all indicators.

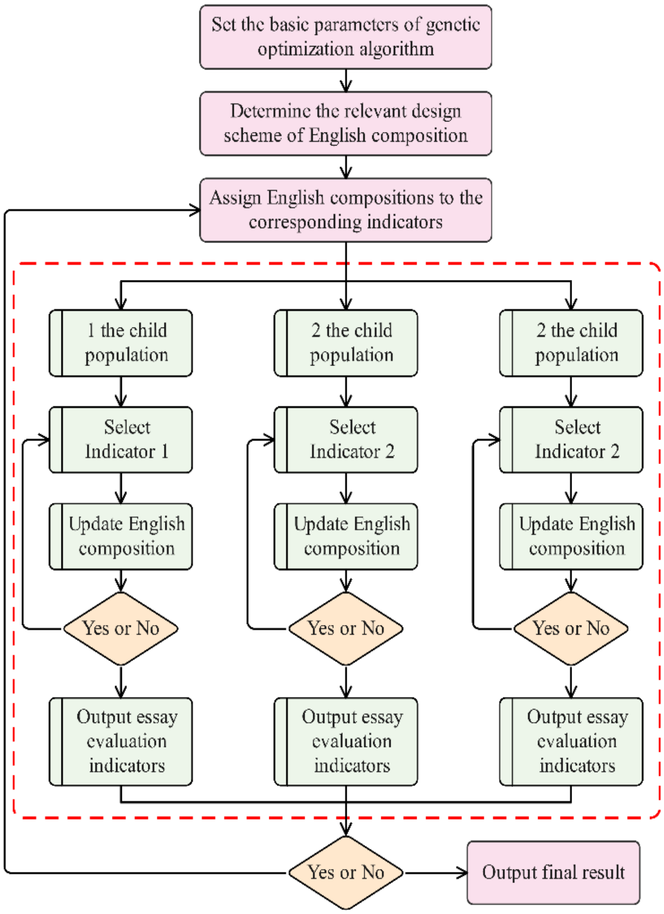

Fig. 10. English composition scoring flow chart based on genetic optimization algorithm model.

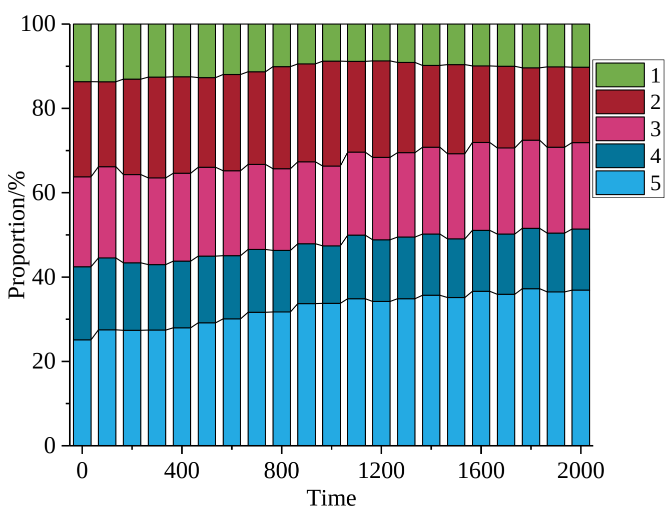

Fig. 11. A summary graph of automatic scoring of English composition based on genetic optimization algorithm and CNN model.

V. Discussion

Genetic optimization algorithm is an optimization algorithm that mimics natural selection and genetic processes, and can find the optimal solution by iteratively optimizing the objective function. In automatic grading of English compositions, genetic optimization algorithms can be used to optimize grading standards and rules to better meet practical teaching needs and English language characteristics. Convolutional neural networks (CNN), as a deep learning model, are suitable for processing data such as images and text. In the automatic grading of English compositions, the CNN model can be used to automatically correct English compositions, achieving automated grading by learning a large amount of English composition sample data. The combination of genetic optimization algorithm and CNN model can first optimize the grading rules of English compositions using genetic optimization algorithm, and then automatically grade English compositions using CNN model. This combined method can fully leverage the advantages of both and improve the effectiveness and accuracy of automatic grading of English compositions.

Neural networks can be used to learn and predict the QoS of a given service combination. Historical data can be used for training, and then this model can be used to predict the QoS of new service combinations. Genetic algorithms can be used to find service combinations that meet certain QoS requirements. This may involve encoding services (such as using genetic coding), and then finding service combinations that meet QoS requirements under the predictive guidance of neural networks [22]. In QoS aware service composition, multiple service quality indicators need to be considered, such as processing time, throughput, latency, etc. By calculating the skyline, it is possible to find the optimal SFC that meets multiple QoS indicators simultaneously in a set of candidate service node combinations. Genetic algorithm is an optimization algorithm that simulates natural selection and genetic processes, suitable for solving the optimal solution of complex problems. In the process of combining service function chains, genetic algorithms can be used to generate initial solutions and perform selection, crossover, and mutation operations on the solutions to continuously optimize the quality of the solutions [20]. Correlation analysis and research show that the genetic optimization algorithm and CNN model can jointly evaluate the teaching effect of English composition automatic scoring better. In addition, relevant data show that different evaluation indexes impact English writing differently. To further verify the influence of the genetic optimization algorithm and CNN model on evaluation indexes of English writing. The validation and prediction curves of the optimization model are drawn, as shown in Fig. 12.

It is seen from the curve changes in the figure, the data on English composition increases linearly at first and then slowly decreases with the number of steps. Model curve presents that the optimization model can well elaborate and represent the changing trend of English composition indicators, so to investigate the impact of automatic English composition scoring on teaching, use the optimization model.

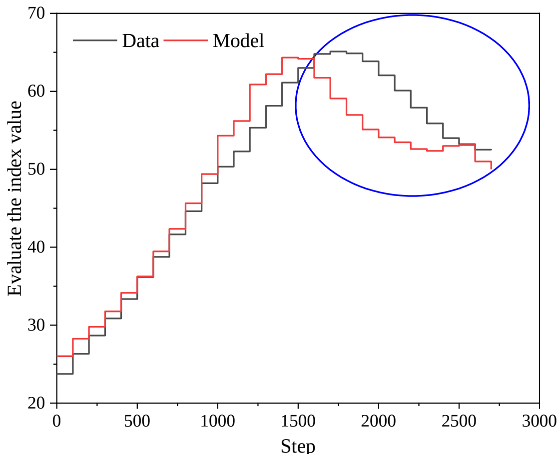

Fig. 12. Model validation and prediction diagrams.

VI. CONCLUSION

1) The real number variation method shows obvious 01 bidirectional variations with the increased iteration number. The corresponding binary variation method displays the linear change of two sections with different slopes. It is worth explaining that the increasing of samples can promote the improvement of test accuracy.

2) The neuron data under the genetic algorithm has typical linear variation characteristics, while the neuron spacing curve under the influence of initial domain radius and time constant presents distinct nonlinear characteristics. This indicates that the original model cannot describe the changes in the test data well, and further optimization and analysis of the model are needed.

3) As the number of iterations increases, the influence curves of the genetic algorithm under four different parameters are shown as follows: Parameter A increases slowly at first and then presents a relatively gentle nonlinear change; Parameter U decreases gradually and then rapidly; Parameter m stays at about 70 on the whole. The parameter I shows typical linear characteristics.

4) The accuracy of the corresponding genetic optimization algorithm can be obtained by calculating the corresponding accuracy index and the accurate total number, and the relevant change curve can be drawn as follows: The curve of accuracy R first showed a significant downward trend and then linearly increased. And the increase in iteration and calculation times can improve the accuracy of calculation results.

However, this article also has certain limitations. The CNN model based on genetic algorithm requires a large amount of computing resources, which will consume significant time during training and usage. Meanwhile, for small image datasets, this may not achieve good performance. This may be difficult to solve specific problems such as accurate positioning of objects in images. The scale and quality of the corpus on which an automatic scoring system relies directly affect the accuracy of the system. Therefore, future research should focus on establishing more complete, standardized, and representative corpora. Expand coverage and classification fineness, improve the universality and accuracy of the system. Natural language processing can only process textual information, while English automatic scoring includes not only text but also information such as voice and images. Therefore, future research should focus on integrating multimodal information to achieve multimodal automatic scoring systems, improving the comprehensiveness and accuracy of scoring.

REFERENCES

[1] D. Peng, G. Tan, K. Fang, L. Chen, P. K. Agyeman, and Y. Zhang, "Multiobjective optimization of an off-road vehicle suspension parameter through a genetic algorithm based on the particle swarm optimization," Math Probl Eng, vol. 2021, pp. 1–14, 2021.

[2] M. Barthwal, A. Dhar, and S. Powar, "The techno-economic and environmental analysis of genetic algorithm (GA) optimized cold thermal energy storage (CTES) for air-conditioning applications," Appl Energy, vol. 283, p. 116253, 2021.

[3] M. Mangera, J. O. Pedro, and A. Panday, "Direct adaptive neural network-based sliding mode control of a high-speed, ultratall building elevator using genetic algorithm," SN Appl Sci, vol. 4, no. 4, p. 100, 2022.

[4] K. Cai, X. Li, and L. H. Zhi, "Extracting time-varying mean component of non-stationary winds utilizing Vondrak filter and genetic algorithm: A wind engineering perspective," International Journal of Structural Stability and Dynamics, vol. 21, no. 11, p. 2150155, 2021.

[5] N. Ben Latifa and T. Aguili, "Optimization of coupled periodic antenna using genetic algorithm with floquet modal analysis and mom-gec," Open Journal of Antennas and Propagation, vol. 10, no. 1, pp. 1–15, 2022.

[6] S. Li et al., "Development of a reduced chemical reaction mechanism for n-pentanol based on combined reduction methods and genetic algorithm," ACS Omega, vol. 6, no. 9, pp. 6448–6459, 2021.

[7] S. Das, T. Samanta, and A. K. Datta, "Improving black tea quality through optimization of withering conditions using artificial neural network and genetic algorithm," J Food Process Preserv, vol. 45, no. 3, p. e15273, 2021.

[8] J. Wang, R. Wang, M. Yang, and D. Xu, "Understanding zinc-doped hydroxyapatite structures using first-principles calculations and convolutional neural network algorithm," J Mater Chem B, vol. 10, no. 8, pp. 1281–1290, 2022.

[9] Y. Zhou, H. Chen, Y. Li, S. Wang, L. Cheng, and J. Li, "3D multi-view tumor detection in automated whole breast ultrasound using deep convolutional neural network," Expert Syst Appl, vol. 168, p. 114410, 2021.

[10] M. AbiarKashani, Y. Alizadeh Vaghasloo, and M. AghaMirsalim, "Optimal design of high-pressure fuel pipe based on vibration response and strength using multi-objective genetic algorithm," Structural and Multidisciplinary Optimization, vol. 64, pp. 935–956, 2021.

[11] D. Żelasko, W. Książek, and P. Pławiak, "Transmission quality classification with use of fusion of neural network and genetic algorithm in Pay&Require multi-agent managed network," Sensors, vol. 21, no. 12, p. 4090, 2021.

[12] S. Güler and S. Yenikaya, "Analysis of shielding effectiveness by optimizing aperture dimensions of arectangular enclosure with genetic algorithm," Turkish Journal of Electrical Engineering and Computer Sciences, vol. 29, no. 2, pp. 1015–1028, 2021.

[13] X. Lin et al., "Optimized neural network based on genetic algorithm to construct hand-foot-and-mouth disease prediction and early-warning model," Int J Environ Res Public Health, vol. 18, no. 6, p. 2959, 2021.

[14] G. Long, Y. Wang, and T. C. Lim, "Optimal parametric design of delayless subband active noise control system based on genetic algorithm optimization," Journal of Vibration and Control, vol. 28, no. 15–16, pp. 1950–1961, 2022.

[15] S. Y. Martowibowo and B. Kemala Damanik, "Optimization of material removal rate and surface roughness of AISI 316L under dry turning process using genetic algorithm," Manufacturing Technology, vol. 21, no. 3, pp. 373–380, 2021.

[16] X. Han, D. Liang, and H. Wang, "An optimization scheduling method of electric vehicle virtual energy storage to track planned output based on multiobjective optimization," Int J Energy Res, vol. 44, no. 11, pp. 8492–8512, 2020.

[17] Z. Ran, W. Ma, C. Liu, and J. Li, "Multi-objective optimization of the cascade parameters of a torque converter based on CFD and a genetic algorithm," Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 235, no. 8, pp. 2311–2323, 2021.

T. Zhu, L. Wang, X. Na, T. Wu, W. Hu, and R. Jiang, "Research on on Novel Fuzzy Control Strategy of Hybrid Electric Vehicles Based on Feature Selection Genetic Algorithm.," Sensors & Materials, vol. 33, 2021.

[19] K. Kassoul, N. Cheikhrouhou, and N. Zufferey, "Buffer allocation design for unreliable production lines using genetic algorithm and finite perturbation analysis," Int J Prod Res, vol. 60, no. 10, pp. 3001–3017, 2022.

[20] T. Zhu, Y. Qin, Y. Xiang, B. Hu, Q. Chen, and W. Peng, "Distantly supervised biomedical relation extraction using piecewise attentive convolutional neural network and reinforcement learning," Journal of the American Medical Informatics Association, vol. 28, no. 12, pp. 2571–2581, 2021.

[21] Y. Ding et al., "AP-CNN: Weakly supervised attention pyramid convolutional neural network for fine-grained visual classification," IEEE Transactions on Image Processing, vol. 30, pp. 2826–2836, 2021.

[22] P. Khosravian, S. Emadi, G. Mirjalily, and B. Zamani, "QoS-aware service composition based on context-free grammar and skyline in service function chaining using genetic algorithm," PeerJ Comput Sci, vol. 7, p. e603, 2021.