Analysis and Application of Antibacterial Drug Resistance Based on Deep Learning

Wei Zhang[1], Yanhua Zhang[2], Qiang Zhang[3], Caixia Xie[4], Yan Chen[5], Jingwei Lei[6*]
Department of Basic Medicine, Zhengzhou Health Vocational College, Zhengzhou 450122, China[1]
Department of Public Education, Zhengzhou Health Vocational College, Zhengzhou 450122, China[2]
Department of Criminal Science and Technology, Henan Police College, Zhengzhou 450046, China[3]
Henan Engineering Research Center for Quality Control and Evaluation of Traditional Chinese Medicine, School of Pharmacy, Henan University of Chinese Medicine, Zhengzhou 450046, China[4, 6]
Department of Clinical Laboratory, Henan Infectious Disease Hospital, Zhengzhou 450052, China[5]

***Abstract*—The continuous improvement of deep learning technology has led to its deeper application in related fields, especially in the detection of antimicrobial resistance in the medical field. In drug resistance detection, the CNN-ATT-TChan model based on the fusion of CNN algorithm and attention mechanism can classify and organize a large amount of antimicrobial resistance data, achieving standardized processing. Based on mature chemical analysis and testing methods, drug resistance test data was obtained, and the training duration and classification accuracy *F* of the model were discussed in combination with the test data. At the same time, based on relevant research literature, the changes in *ROC* curves and *AUC* values between different models were compared. The results showed that the CNN algorithm using fusion attention mechanism can improve the training time of the model and also improve the classification accuracy of the model. Therefore, the application of CNN-ATT-TChan model combined with attention mechanism in the detection of antimicrobial resistance provides more support for the development of antimicrobial resistance testing.**

Keywords*—Deep learning; antibacterial drug; CCN algorithm; pharmaceutical chemical analysis; drug resistance testing

I. INTRODUCTION

The development of antibiotics, especially antibiotics, has greatly reduced the incidence rate and mortality of many infectious diseases. However, with the continuous use of antibiotics, bacteria can develop resistance to antibiotics, which has become one of the main challenges threatening global public health. Antibiotic resistance refers to the ability of bacteria to resist or become tolerant to chemotherapy agents, antimicrobial agents, or antibiotics. This tolerance may be achieved through gene mutations or exogenous DNA in the transmission of R-factors. Bacteria that are resistant to multiple antibiotics are also known as multidrug-resistant, while bacteria that are considered to be extensively or completely resistant are referred to as super bacteria [1-2]. Resistance generally occurs when many infectious bacteria gradually adapt to first-line or even second-line antibiotics and develop bacterial resistance in the context of antibiotic abuse and overuse. This type of problem poses a huge burden and impact on medical hygiene, veterinarians, agriculture, and society. Affected by drug resistance, treatment plans may be limited or unavailable, posing a higher risk of death, longer hospital stays, and recovery processes for patients, sometimes even leading to long-term incapacity. Therefore, it is necessary to fully analyze the internal reasons for the formation and development of antimicrobial resistance, and based on this, study the role of antimicrobial resistance as shown in Fig. 1. The process of antibiotic resistance occurrence is presented from the perspective of drug resistance gene transfer.

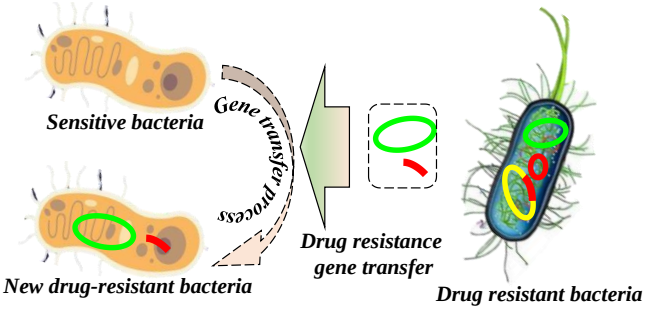


Fig. 1. Process of antibiotic resistance occurrence.

Antimicrobial resistance analysis is a complex process, and deep learning, as one of the key core technologies of big data analysis, provides more means for the improvement and development of antimicrobial resistance analysis. At present, the basic data in the medical field is complex, and electronic medical records, diagnosis and treatment data, laboratory and inspection data, as well as various management data, have become major issues faced by drug resistance analysis. Therefore, further improving the intelligent analysis and processing technology of antimicrobial resistance in the medical field has become a key link in drug detection in this field [3-4]. In the medical field, although there are already many chemical testing methods such as drug resistance test kits and analysis software being applied, with the continuous improvement of data utilization and analysis requirements, relying solely on traditional chemical testing methods cannot provide better support for drug resistance analysis in the in the medical field. The use of antibacterial drugs is mainly determined based on experience and habits, resulting in a lack of comprehensive analysis of the indications and contraindications of antibacterial drugs, as well as the personal characteristics of patients, and the inability to achieve the best expected treatment effect.

In the medical field, traditional bacterial culture methods are often used for drug resistance testing, which takes at least 3-5 days from identifying bacteria to completing drug sensitivity testing [5]. This traditional testing method delays the optimal time for doctors to analyze patients' antimicrobial resistance, posing challenges and difficulties for clinical drug use. At present, there is relatively little research on data analysis and processing of antibacterial drugs based on deep learning algorithms. Researchers have used the XGBoost algorithm of machine learning to construct MIC prediction models between the genome of non-Salmonella typhimurium and multiple antibiotics. Some scholars have also conducted exploratory research on the rational use of antibiotics using machine learning immune genetic algorithms and deep learning short-term memory network models [6-8]. Based on deep learning algorithms and combined with modern chemical methods, it provides a more accurate and efficient means for the analysis of antimicrobial resistance in the medical field. At the same time, it can significantly shorten the testing cycle and provide testing results in a timelier manner [9-10]. Chemical methods play an important role in drug analysis and research and development. Through their diverse analytical methods, from molecular structure to compound labeling, from component testing to local drug analysis, this method provides an effective way to solve drug analysis problems. In the field of medicine, the updates in deep learning technology and the development of chemical methods have provided more comprehensive analytical methods for the analysis of antimicrobial resistance.

This article combines the convolutional neural network (CNN) algorithm in deep learning and proposes a dual channel CNN terminal model in the medical field that integrates attention mechanisms based on the analysis and summary of completed drug resistance experimental data. Using a CNN model for feature extraction and completing self-learning, assigning different weights to information of different importance, and constructing a classification model between testing samples and antibacterial drugs. Using this model to determine the tolerance level of the test samples to antibiotics can provide auxiliary effects for the rational use of antibiotics in the medical field. At the same time, doctors in Harbin can provide the most suitable decision support for individual patients in clinical medication.

II. Analysis of Drug Resistance in the Context of Deep Learning

A. Analysis of Deep Learning Algorithms

Deep learning is an important branch of machine learning. By learning deep nonlinear network structures, deep learning trains features layer by layer, gradually transforming the feature representation of samples in the original space into a new feature space, demonstrating the powerful ability to learn the essential features of datasets from the sample set. The biggest advantage of deep learning is that it can automatically learn features and achieve the abstraction process of features. It adopts multi-layer complex structures or multiple layers composed of multiple nonlinear transformations for data processing. The use of deep learning for drug resistance testing and analysis in the medical field has unique advantages. It can fully explore the deep relationships hidden in a large amount of medical data, obtain more abstract feature values through learning neural networks, and improve the feasibility and accuracy of antimicrobial resistance analysis in the medical field [11-12]. As shown in Fig. 2, a schematic diagram of the network structure of the deep learning algorithm is provided. The algorithm consists of three dimensions: input layer, hidden layer, and output layer, with relevant nodes processing and analyzing data in each layer.

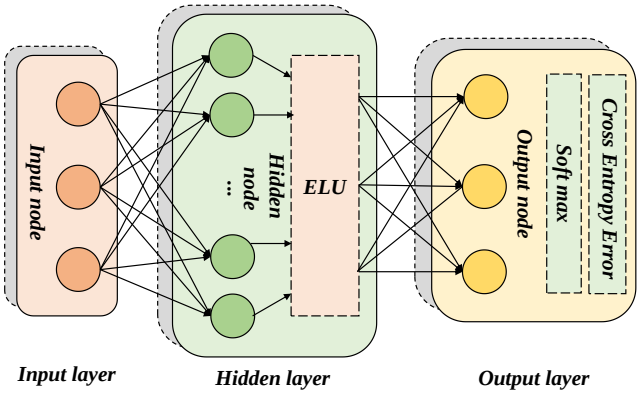


Fig. 2. Schematic diagram of deep learning algorithm network structure.

B. Application Status of Deep Learning Algorithms in Drug Resistance Testing

The continuous development of deep learning algorithms has led to the application of more and more technologies in the medical field, promoting the continuous updating and progress of drug resistance testing methods in the medical field. At the same time, thanks to the application of deep learning technology in drug research and development, people have also made many beneficial explorations and attempts [13-15]. The study of gene expression and protein structure, compound screening, and the design and analysis of clinical trials for drug design are all influenced by artificial technology. In addition, the application of various digital devices in clinical trials and daily patient monitoring will also benefit from the help of artificial intelligence technology. Driven by machine learning and deep learning technologies, significant progress has been made in the analysis of antimicrobial resistance testing in the medical field. By using CNN technology in deep learning, testers can obtain more detailed data on antimicrobial resistance testing and analysis, and conduct more effective research and analysis on drug resistance, Furthermore, more accurate drug resistance diagnosis and prediction can be made for patients [16]. The application of deep learning algorithms in the testing of antimicrobial resistance can help evaluate drug resistance in a shorter time with fewer cases and reduce the cost of clinical trials. As shown in Fig. 3, a schematic diagram of the application of the CNN algorithm model in drug resistance testing and analysis is provided.

Based on deep learning algorithms, comprehensive analysis of a large amount of antimicrobial resistance testing data has become a trend in the medical field. In traditional medical practices, doctors also use antibiotics based on their

own experience and habits, lacking a more systematic and comprehensive consideration of the indications and contraindications of antibiotics, as well as the personal characteristics of patients [17]. A deep learning based identification model for antibacterial drug applications. Firstly, the coding distribution representation of the antibacterial drug is obtained. Then, a CNN model is trained for feature extraction and self-learning. Finally, a reasonable rule and feature library for single disease and complications is generated. Therefore, using the CNN model in deep learning algorithms can more efficiently detect and analyze antimicrobial resistance [18].

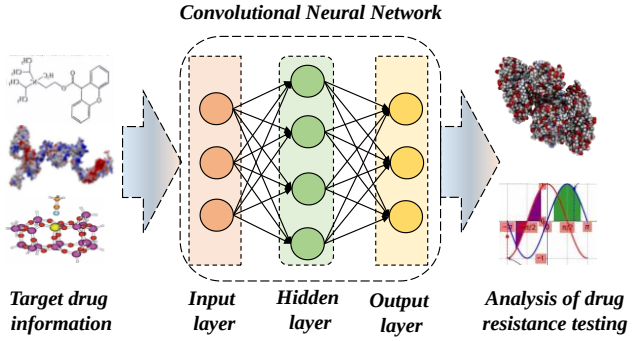


Fig. 3. Application of CNN algorithm in drug resistance testing and analysis.

III. Construction of a Drug Testing Model Based on CNN Algorithm

A. CNN Algorithm Theory

CNNs are mainly used in fields such as image processing and recognition. They are a multi-layer perception mechanism mainly used to process two-dimensional images. As shown in Fig. 4, a schematic diagram of the CNN module in deep learning algorithms is provided. The basic structure of CNN consists of an input layer, a convolutional layer, a pooling layer, a fully connected layer, and an output layer. The pooling layer merges adjacent multiple nodes into a similar feature to reduce the amount of training data, extract hidden features from the project, conduct multi-layer training, and perform dimensionality reduction processing. Finally, the CNN is trained as a whole [19-20].

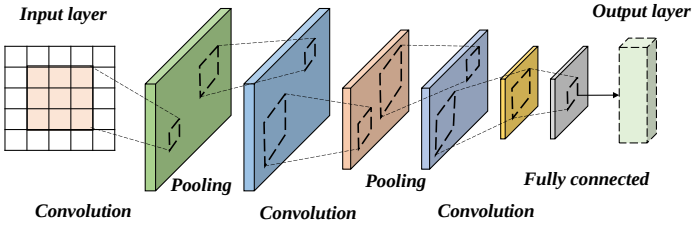


Fig. 4. Schematic diagram of CNN model.

In the CNN model algorithm, L represents the number of layers of the neural network, i represents the label of a certain layer, then $w_{h,w}^{i}$ represents the connection weight from the wth neuron in layer i-1 to the h-th neuron in layer i, b_h^i represents the bias of the h-th neuron in layer i, and Z_h^i represents the weighted input of the h-th neuron in layer l. The weighted input of the h-th neuron in the l-th layer can be represented by Formula (1):

$$Z_h^i = \sum_h w_{h,w}^i a_w^{i-1} + b_h^i \tag{1}$$

For convolutional layers, their main function is to extract features, perform convolution calculations on the vectorized text vector and the convolution sum, and perform operations on extracting word information before and after. The obtained convolution results are processed through activation function operations, and the convolution function results are used as the output results of this layer. The convolutional kernel is represented by w, where $w \in R^{hk}$. Where h represents the height of the convolutional kernel window, and k represents the dimensional size of the word vector. W_{ih} represents a sequence of words with a length of h (W_i, W_{i+1}, …, W_{i+h}), Wi represents a word, and the value of each eigenvalue can be calculated using Formula (2):

$$y_i = f\left(\sum_{h=1}^{n}\sum_{w=1}^{n}\sum_{d=1}^{m} x_{h,w,d} \times w_{h,w,d}^i + b^i\right) \tag{2}$$

Among them, n represents the height and width of the convolutional kernel, m represents the depth of the convolutional kernel, $x_{h,w,d}$ represents the value at the region connected to the convolutional kernel (h, w, d), $w_{h,w,d}^i$ represents the weight parameter at the convolutional kernel (h, w, d), b^i represents the bias of the convolutional kernel, and $f(\cdot)$ represents the activation function. In addition, the input layer x is convolved to obtain the feature vector y, which can be represented by Formula (3):

$$y = (y_1, y_2, y_3, \ldots, y_{i-h+1}) \tag{3}$$

In addition, the pooling layer can reduce the dimensionality of the extracted feature information, reducing the size of the feature map, simplifying the computational complexity of the network, and avoiding overfitting to a certain extent. On the one hand, feature compression is performed to extract the main features. This layer is mainly used to reduce the dimensionality of the feature map obtained by convolution, and the maximum pooling method shown in Formula (4) is often used to extract local optimal values:

$$Y = max(y_i) \tag{4}$$

The fully connected layer in a CNN is usually located after several convolutional and pooling layers, and is also the last few layers of the CNN. Its essence is a multi-layer perceptron model, where each neuron in the fully connected layer is connected to all neurons in the previous layer. The most direct function of the fully connected layer is to compress the two-dimensional feature map into a one-dimensional feature vector, in order to connect with the final output layer of the network to output the final classification results [21]. At the same time, it can also to some extent integrate local information with class differentiation in convolutional and pooling layers. The role of the previous convolutional and

pooling layers is to map a shallow feature into a deep feature space, while the fully connected layer maps these deep feature representations into the classification space of the class. Therefore, the fully connected layer also has a certain classification effect.

The use of nonlinear factors can better solve complex problems, and activation functions can help introduce nonlinear factors, adding some nonlinear factors to neural networks, and enabling neural networks to better solve more complex problems. Common saturated linear functions include *Sigmoid*, *Tanh* function, and *Tanh* activation function, which converge faster than *Sigmoid*, but there is still a phenomenon of gradient vanishing during training [22-23]. As shown in Formulas (5) and (6), the activation function expression is given:

$$Sigmoidf(x) = \frac{1}{1+e^{-x}} \tag{5}$$

$$Tanhf(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{6}$$

In this article, the specific method is first based on deep learning methods. The input part involves the name, function, dosage, indication, contraindications, etc. of antibacterial drugs. The word vector tool is used to train the word vector, and then the transformed sentence is input into a CNN for convolution and pooling. Then, the fully connected deep learning features are obtained.

B. Construction of Drug Testing Model Based on CNN Algorithm

1) Text classification: In the medical field, text classification of data is required when using CNN algorithm for antibiotic resistance testing and analysis. Text vector is a higher-level expression of text, which can be used as a partial feature vector for antibacterial drug text data classification. After training and learning through standard neural networks, vector V is obtained as another basis for antibacterial drug text data classification [24]. The weights of feature vectors in classification are α and (1-α). The probability calculation formula for text belonging to a certain type of effect is shown in equation (7):

$$P = softmax\left[\alpha \times W_c C + (1-\alpha) \times W_c H + b_c\right] \tag{7}$$

Using the negative logarithmic likelihood of the correct label as the training loss, the calculation formula is shown in formula (8):

$$L = -\sum_{d} log_a x d_j \tag{8}$$

Where j is the label of text d, $R = 1 - \frac{C}{A+C}$ represents the probability that text d belongs to effect j.

2) Modeling of drug sensitivity testing data: The dual channel CNN model proposed in this article integrates attention mechanism. After preprocessing the original test data, it is vectorized and modeled. As input data, it is sent to the CNN of two independent and different depth channels. After each channel undergoes several convolutional and pooling operations, attention mechanism is introduced. By fitting multiple sets of weight vectors to characterize the importance of each feature component, the feature data of the two channels is fused, and finally, the softmax function is used to achieve classification output. The data source studied in this article is bacterial drug sensitivity testing data. Each test sample includes the patient's age, gender, department, submission date, sample type, bacterial species, testing chemical reagents, testing report date, and other submission data, as well as the patient's minimum inhibitory concentration (MIC) testing results for each antibacterial drug. With the development of physical, chemical, and molecular biology technologies, various new bacterial resistance testing technologies have emerged. As shown in Table I, common chemical methods and their characteristics for detecting antimicrobial resistance are presented.

TABLE I. CHEMICAL METHODS AND CHARACTERISTICS OF ANTIMICROBIAL RESISTANCE TESTING [25-26]

Method	Characteristics	Application Scenario
Nucleic acid hybridization technology	Short testing time, time-saving and fast, strong specificity, and high sensitivity	Rapid diagnosis of pathogens in clinical settings
PCR	Short testing time, time-saving and fast	Rapid diagnosis of pathogens in clinical settings
Disk diffusion	Simple, repeatable, without the need for expensive device support	Qualitative testing of bacterial sensitivity to antibiotics
Testing method based on time-of-flight mass spectrometry technology	Microbial cultivation takes a long time and requires exploration of antibiotic usage conditions	Rapid testing of drug resistance

Select multiple attributes from the submitted data that may affect the tolerance value of antibacterial drugs as input features, and use the testing result MIC as a classification label. Considering that different bacterial strains have different sensitivities to the same antibacterial drug, and the sensitivity values of the same bacterial strain to the same antibacterial drug vary among different specimens [27]. Therefore, it is necessary to establish a classification model for each antibacterial drug, and output its MIC classification value for the current antibacterial drug based on the input multi feature test data. The vectorized representation of sample data containing multiple features is an important prerequisite for feature extraction and fusion using CNNs. This article draws inspiration from the approach of constructing a word vector model in text problems, where each attribute feature value corresponds to a word and is represented by a vector. The dimension of the feature vector is the number of all feature values. As the feature vector dimension of the drug sensitivity data selected in this article is 48, there is no problem of dimensional disaster, and the correlation between various features is relatively weak. Therefore, the traditional *One-hot* method is used to construct the feature vector model. A sample

data containing multiple features corresponds to a sentence in the text problem, which is modeled as a two-dimensional matrix, and each row of the matrix is a feature vector. As shown in Fig. 5, a dual channel convolutional neural network model structure diagram integrating attention mechanism is presented.

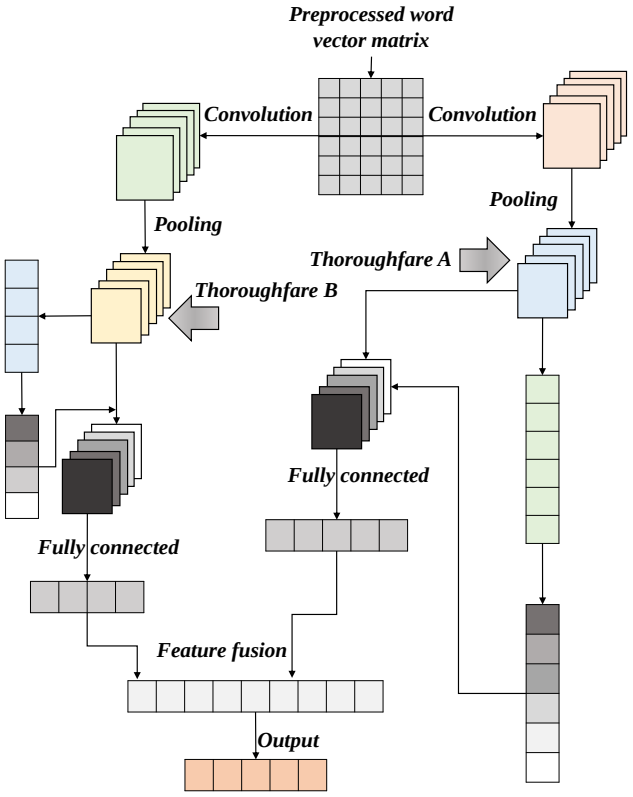


Fig. 5. A dual channel CNN model integrating attention mechanism.

The features extracted by CNNs are often related to the depth of the network. The deeper the network level, the easier it is to extract abstract features that represent the whole. The shallower the network level, the easier it is to extract detailed features that represent local regions. In view of this, this article designs two types of dual channel CNNs with depths of six and four layers, respectively, to extract features at different levels of abstraction in deep and shallow layers. In addition, considering that the feature matrix size of the testing sample is not large, pooling dimensionality reduction is weakened. The upper channel network consists of five convolutional layers with different kernel sizes and numbers, three maximum pooling layers with the same step size, and 1 fully connected layer, the down channel network consists of three convolutional layers with different kernel sizes and numbers, one maximum pooling layer, and one fully connected layer.

3) Attention mechanism and model training: The implementation of the Attention Mechanism is achieved by preserving the intermediate output results of the CNN encoder on the input sequence, and then training a model to selectively learn these inputs and associate the output sequence with them when the model outputs [28-29]. Although the model may increase computational complexity after using this mechanism, its performance level can be improved. In addition, using this mechanism helps to analyze the impact of relevant information on the final generated sequence.

The attention mechanism assigns appropriate weights to each feature component, selects important features, and focuses on this information. It is divided into three parts, namely Squeeze, Excitation, and Attention. Assuming that the feature information after convolution and pooling is a two-dimensional matrix containing n feature vectors, denoted as $U_{m\times n}$, after squeezing, the vector Z is obtained. The calculation of the squeezing function (F_{sq}) is shown in Formula (9):

$$Z = \sum_{i=1}^{m} U_i \Big/ m \tag{9}$$

The vector Z is stimulated by Formula (10) to obtain the attention weight vector a. Multiply the weight vector a with the feature matrix U to generate a feature output U' with attention mechanism, and $U'=aU$.

$$a = softmax\left[w_s \times tanh(w_z \times z)\right] \tag{10}$$

where, m is the dimension of the feature vector, U_i is the *i-th* feature vector, *softmax* (·) and *tanh*(·) are activation functions, and w_s and w_z are learnable weight matrices.

In addition, during the model training process in the CNN algorithm, the *softmax* function is used to classify the feature vectors after dual channel fusion, and the predicted probability value $P(k|x)$ of the output result category is shown in Formula (11):

$$P(k|x) = e^{x|k|} \Big/ \sum_{i=1}^{n} e^{x|i|} \tag{11}$$

The cross-entropy loss function is used to measure the difference between the predicted probability and the real label, and the calculation is shown in Formula (12). This is used for backpropagation training of the model, with the goal of minimizing the loss of the function [30]:

$$L = \sum_{i=1}^{n} q(k|x) log\left[P(k|x)\right] \tag{12}$$

where, $q(k|x)$ represents the actual encoding vector of the corresponding category k of the current sample, which is the real label.

IV. Analysis of Drug Resistance Testing and Model Application

A. Drug Resistance Testing

1) Source of strains and testing instruments: When using a deep learning CNN algorithm to detect antimicrobial resistance, the relevant data comes from bacterial drug sensitivity testing data in a tertiary hospital's intensive care

unit. 120 strains of Acinetobacter baumannii were isolated from clinical specimens, and duplicate strains from the same patient and site were discarded. The standard strains are Escherichia coli ATCC25922 and Pseudomonas aeruginosa ATCC27853, using drugs such as Ampicillin/Sulbactam (AMP/SCF), Fefepime (FEP), Cefoperazone/Sulbactam (CSL/SCF), Cefotaxime (CTX), Ceftazidime (CAZ), Ceftriaxone (CRO), Ciprofloxacin (CIP), Gentamicin (GEN), Imipenem (IMP), Meropenem (MEM), Piperacillin/Tazobactam (TZP). Piperacillin (PIP), Levofloxacin (LVX), Amikacin (AMK), Compound Sulfamethoxazole (SXT). In addition, the culture medium is Mueller Hunton (M-H) agar and broth. As shown in Table II, the relevant equipment for drug resistance testing is provided.

TABLE II. DRUG RESISTANCE TESTING EQUIPMENT AND INSTRUMENTS

Types	Device
Main instruments	Ultra low temperature water tank
	A400 multi-point vaccination device
	Electro-heating standing-temperature cultivator
	VITEK Microbial Fully Automatic Identification Instrument
	Paper dispenser
	JY2002 Electronic Balance
	ZHP-2102 Intelligent Constant Temperature Shaker Incubator
Other auxiliary instruments	Pipette, glass test tube, PCR tube, EP tube, bacterial culture dish, etc

2) Drug resistance testing methods: The sensitivity of antibacterial drugs was determined using K-B agar diffusion method and M-H agar dilution method. The screening criteria for multiple drug resistance were: Acinetobacter Baumanii against cephalosporins, aminoglycosides, carbapenems, and enzyme inhibitors β Individuals with resistance to three or more antibiotics, including lactams and fluoroquinolones, are identified as multidrug-resistant Acinetobacter baumannii.

For the preparation of flat plates, it is necessary to mark the different concentrations on the plates, dilute the antibiotic stock solution, and then pour the antibiotics into the plates using a double dilution method. Wait for M-H agar to cool to about 50℃, and pour a certain amount of antibacterial drugs into the plate with the calculated final concentration. Inoculate the quality control strain and the test strain onto an M-H agar plate, incubate overnight at 37℃, and then select 4~5 colonies to be cultured overnight in 1ml of M-H broth. Dilute 10 times, which is equivalent to 0.5 Mach's turbidity. Dip the bacterial solution with a multi-point inoculum and inoculate it onto an antibacterial drug plate. The amount of inoculation per point is approximately 10^4~10^5CFU. Place the inoculated antibacterial drug plate in an incubator at 35℃ for 20~24 hours and observe the results. Each operation includes quality control strains.

In addition, for the drug sensitivity testing process, the antibacterial *K-B* method is used to determine the size of the antibacterial zone. The fresh pure bacterial solution with a concentration of 0.5 microns is uniformly coated with sterile cotton swabs on a 4mm thick and 90mm diameter M-H agar drug sensitive plate. Each plate is coated with six pieces of drug sensitive paper, and incubated at 35℃ for 16~18 hours to measure the size of the antibacterial zone on the plate and interpret the results. The results were determined to be sensitive, medium sensitive, or resistant according to the relevant regulations regarding the size of the antibacterial zone.

B. Application and Analysis of Testing Models

Based on the above statistics of antimicrobial resistance testing results, a total of 12765 related data were completed. The actual production data is quite complex and must be preprocessed before being used in experiments. The data preprocessing work in this article includes data filtering and screening, outlier testing, and standardization. Firstly, data filtering involves screening raw data, removing irrelevant indicators, and selecting six attribute features that may affect antimicrobial resistance values as inputs. The corresponding feature inputs are patient gender, patient age, source department, specimen type, bacterial species, and related chemical detection methods. Secondly, outlier testing is mainly based on expert testing experience to identify unreasonable testing results in the original data. At the same time, testing samples with more missing values and fewer records are also included in the outlier data. The final dataset used for the experiment consists of 3562 pieces. Then there is the standardization of data processing, mainly for the convenience of later data processing, quantifying the five features using numerical values to achieve unified quantization methods.

1) Evaluating indicators: In the process of detecting antimicrobial resistance, this article uses the F metric value as the evaluation criterion for the classification results. The process of classifying the test text d in this method is as follows: first, calculate the text similarity between d and each text in the training sample set. The similarity is generally measured using Euclidean distance, and based on the text similarity, find the most similar K closest training texts. The K sample categories are used as candidate categories for d, Then, based on the similarity between the text to be classified d and these K neighbors, the weight of the K nearest neighbor text category is used. The sum of the class weights of the neighboring texts in each category is used as the similarity between the category and the test text. Then, the text to be classified d is divided into the category with the highest weight. The calculation process of text similarity between the text d and di to be classified is given by Formula (13):

$$Sim(d_i,d)=\frac{1}{d-d_i}=\frac{1}{\sqrt{\sum_{j=1}^{n}(x_{i,j}-x_j)^2}} \tag{13}$$

For the weight of the sample d to be classified and the c_j of each class, Formula (14) can be used to calculate:

$$p(d,c_j)=\sum_{i=1}^{n} sim(d_i,d)y(d_i,c_j) \quad (14)$$

where, $y(d_i,c_j)$ is a class attribute function, and when $d_i \in c_j$, there exists $y(d_i, c_j)=1$. When $d_i \notin c_j$, $y(d_i, c_j)=0$.

In addition, this method takes into account both precision (P) and recall (R) indicators, with P, R, and F calculated by Formula (15):

$$\begin{cases} P = \dfrac{TP}{TP+FP} \\ R = \dfrac{TP}{TP+FN} \\ F = \dfrac{2\times P\times R}{P+R} \end{cases} \quad (15)$$

where, TP represents the number of samples divided into positive classes, FP represents the number of samples divided into positive classes for negative classes, FN represents the number of samples divided into negative classes for positive classes, $TP+FP$ represents the actual number of samples classified, and $TP+FN$ represents the expected number of samples.

To further analyze the accuracy of antimicrobial resistance testing, Formula (16) can be used to determine the accuracy of the test results:

$$Accuracy=\frac{TP+TN}{TP+FP+TN+FN} \quad (16)$$

2) Experimental environment and parameters: To ensure the objectivity and effectiveness of antimicrobial resistance testing results, a ten fold cross validation method was adopted in the experiment. The dataset was divided into te parts, with 1 part being the test set and the remaining nine parts being the training set. The 10 samples were randomly rotated for 10 experiments, and the average value was taken as the final result of the model. After multiple experiments, a set of optimal model parameters was determined, and the relevant values are shown in Table III.

TABLE III. SETTING OF MAIN PARAMETERS OF THE MODEL

Parameters	Parameter Value
Number of convolutional kernels in the upper channel (channel A)	64,128,128,128, 128
Number of convolutional kernels in the lower channel (channel B)	64,128,128
Upper channel convolution kernel size (channel A)	2,3,3,3,2
Lower channel convolution kernel size (channel B)	2,3,2
Output layer ratio	0.5
Learning rate	0.001
Maximum number of iterations	500

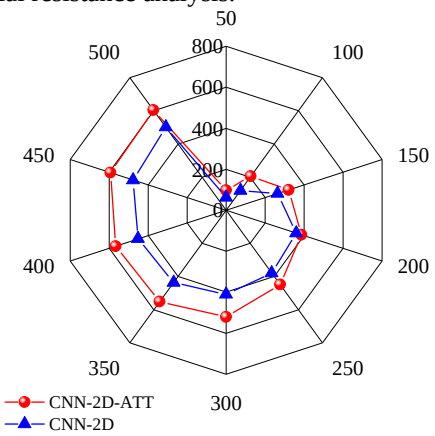
To better evaluate the classification performance of the dual channel CNN model (CNN-ATT-TChan) proposed in this article, which integrates attention mechanisms, in antimicrobial resistance testing data, comparative experiments were designed between different models. As shown in Table IV, the types and basic information of the models used for comparison are provided.

TABLE IV. COMPARISON OF DIFFERENT TYPES OF MODELS

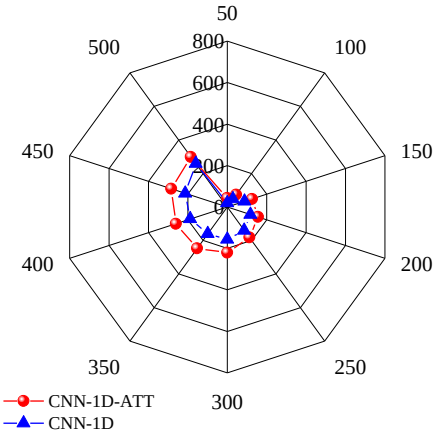
Models	Description
CNN-2D	The convolution kernel of this model is two-dimensional, including multi-layer convolution and pooling operations.
CNN-2D-ATT	This model adds an attention mechanism based on feature components on the basis of CNN-2D.
CNN-2D-TChan	This model is based on CNN-2D and designed with two upper and lower channels, which perform multiple convolution and pooling operations for feature fusion and achieve classification output.
CNN-1D-MChan	This model adopts multiple one-dimensional convolutional kernels with different widths but the same length, each with a length equal to the feature vector dimension of drug sensitivity data.
CNN-1D-Mchan-NP	This model is based on CNN-1D-MChan and removes pooling operations. After each channel undergoes one-dimensional convolution, it directly performs vector fusion and enters the fully connected layer.
ResNet-18	The basic architecture of this model is *ResNet*, with a depth of 18 layers (referring to the weight layer of the network), excluding the batch normalization layer and pooling layer
RF	The random forest algorithm model is a mainstream ensemble learning method used for classification and regression.

3) Analysis of model testing results: Based on the sorting and analysis of the above basic models, in order to verify the impact of adding attention mechanisms in one-dimensional and two-dimensional convolution models on model performance, as shown in Fig. 6, the impact of attention mechanisms on model training duration in one-dimensional and two-dimensional convolution cases is presented. As shown in the figure, for a two-dimensional convolutional model, with the assistance of spatial domain attention mechanism and mixed attention mechanism (spatial+channel), the main role of attention mechanism in one-dimensional convolution and two-dimensional convolution is to significantly increase the training time of the model. For the two-dimensional convolutional model, with the introduction of attention mechanism, when Epochs is 500, the two are 604s and 501s respectively, with a difference of 103s, which is about 20.5% higher overall. In the one-dimensional convolutional model, after introducing attention mechanism, the training duration of the two is 298s and 258s respectively (when Epochs is 500), with a difference of 15.5% between the two. Therefore, it can be seen that when attention mechanism is introduced, the training duration of both will increase to varying degrees. After adding attention mechanism, the training duration has been improved to a certain extent. The training duration of one-dimensional convolutional network

(CNN-1D-ATT) with added attention mechanism is relatively small compared to one-dimensional convolutional network (CNN-1D) without added attention mechanism. The two-dimensional convolutional network (CNN-2D-ATT) with added attention mechanism has significantly improved and improved training time compared to the two-dimensional convolutional network without added attention mechanism (CNN-2D), which also lays the foundation for the accuracy of antimicrobial resistance analysis.



a) Comparison between CNN-2D-ATT and CNN-2D models



b) Comparison between CNN-1D-ATT and CNN-1D models

Fig. 6. Impact of introducing attention mechanism into 1D/2D models on training duration.

Based on the analysis of the training duration of relevant models, in order to further analyze the impact of attention mechanism introduction on the accuracy of different models, as shown in Fig. 7, the trend of F values of one-dimensional convolutional models and two-dimensional convolutional models under different Epochs is presented. As shown in the figure, for the two-dimensional CNN model, with the introduction of attention mechanism, the classification accuracy of CNN-2D-ATT is significantly higher than that of CNN-2D. For example, when Epochs is 500, the F values of the two are 76.1% and 66.4%, respectively, with an overall difference of 9.7%. At the same time, comparing one-dimensional CNN models that also introduce attention training mechanism with two-dimensional CNN models, it can be found that when Epochs is 500, the former has an F-value of 57.2%, while the latter has an F-value of 76.1%, an improvement of about 18.9%, and the improvement effect is significant. The main reason for the improvement and enhancement of this effect is, on the one hand, due to the small number of layers in the one-dimensional convolutional network model, resulting in a decrease in the overall complexity of the network model. On the other hand, the reduction in the size of the convolutional kernel and output feature during the one-dimensional convolutional process leads to a decrease in the time complexity of individual convolutional layers.

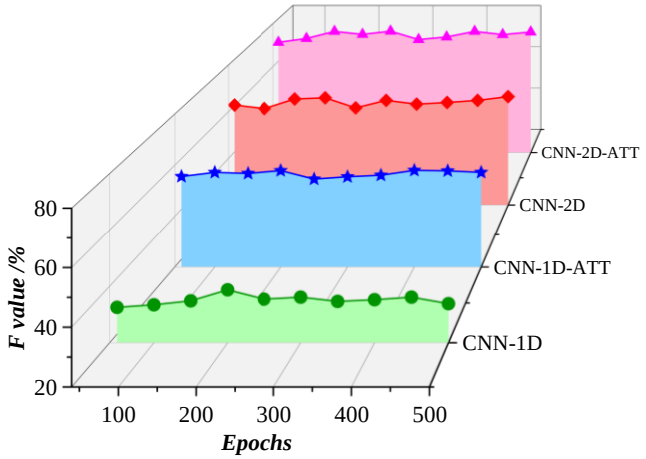


Fig. 7. The influence of attention mechanism on F-value in 1D/2D models.

In addition, in order to compare the overall performance of the model constructed in this article (CNN-ATT-TChan) with other models, based on the analysis of the overall performance of different models, as shown in Fig. 8 and Fig. 9, the F-values of different models and the changes in training time of different models were presented. From Fig. 8, it can be seen that in the comparison of the F values of the above models, the maximum value is the CNN-ATT-TChan model, which is about 72.4%, indicating that the model has better overall performance and can demonstrate better classification accuracy. The model with the lowest F-value is the CNN-1D Mchan model, which is approximately 50.6%. In addition, in Fig. 7, the comparison of the training time of different models with Epoch of 50 is shown. The average time of the above models is about 166 seconds, and the ResNet-18 model with the longest required time is about 653 seconds. The training duration of the CNN-ATT-TChan model is approximately 103 seconds, which is a decrease of 63 seconds compared to the average. Therefore, the CNN-ATT-TChan model constructed in this paper exhibits good classification performance when the experimental data for antimicrobial resistance testing is large and the amount of feature engineering is increased.

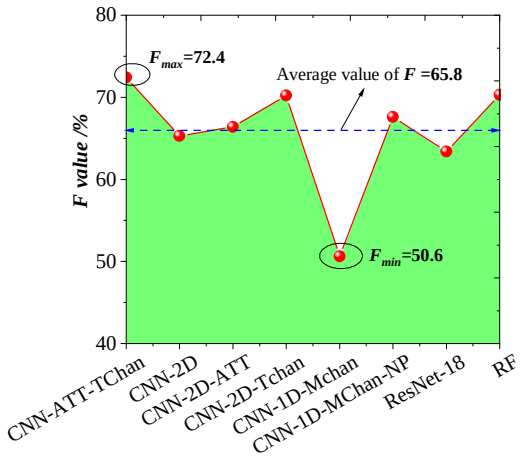


Fig. 8. Comparison of F-values in different calculation models.

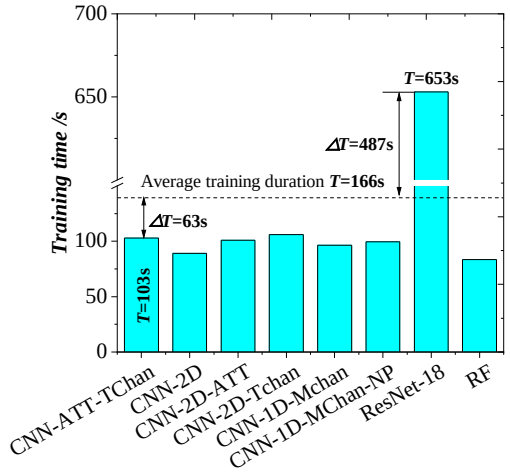


Fig. 9. Comparison of training time for different models.

As shown in Fig. 10, a comparison of the ROC curves between this model and other model is presented. From the figure, it can be seen that the AUC corresponding to the ROC curve in this paper is 0.86, while the AUC of the VGGC16+LSTM model in study [31] is about 0.71. Therefore, the model in this article has better detection performance.

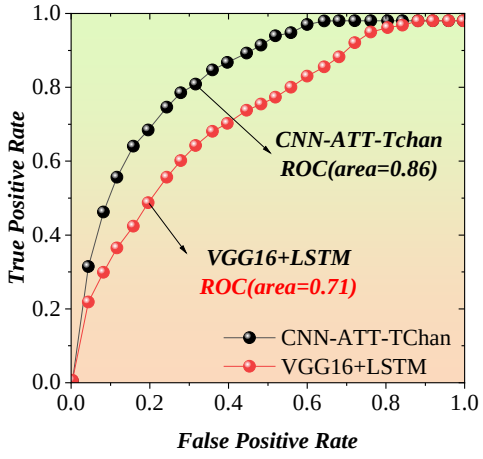


Fig. 10. Average ROC curves of different models.

Through the comparison of the above models and related indicators, it can be found that the CNN ATT Tchan model constructed in this article can effectively improve the problems of low detection accuracy and relatively long detection time in traditional models in drug detection. While solving the problem of low detection accuracy, it achieves the goal of shortening the detection cycle of antibiotic resistance. In this article, improvements have been made in model detection efficiency and detection cycle performance, but further research on drug resistance detection is still a worthwhile direction. For example, deeper research can be conducted by combining a large amount of clinical data, providing more support for drug resistance testing work.

V. Conclusions

The continuous development and application of CNN algorithm in deep learning technology have promoted the continuous progress of antimicrobial resistance testing technology in the medical field. This article constructs a CNN-ATT-TChan model based on the application of CNN algorithm in drug resistance testing, and compares and analyzes the performance of this model in antimicrobial drug resistance testing. The changes in the training time and classification accuracy of the model are presented. The main conclusions are as follows:

By combining the CNN algorithm and attention mechanism, the constructed CNN-ATT-TChan antibacterial drug resistance testing and analysis model can achieve better overall testing performance compared to other models. It can effectively improve important parameters such as training time and classification accuracy in the antibacterial drug resistance testing process in the medical field, and achieve improved testing performance.

By integrating attention mechanism into one-dimensional CNN and two-dimensional CNN models, for two-dimensional CNN models, with the introduction of attention mechanism, when Epochs is 500, the training duration is increased by about 20.5%. In the one-dimensional CNN model, the introduction of attention mechanism resulted in a 15.5% increase in training time.

The constructed CNN-ATT-TChan antibacterial drug resistance testing model exhibits better overall classification accuracy and small sample classification, with certain feasibility and effectiveness. Compared to other models, the classification accuracy can be improved by up to 21.8%.

Acknowledgment

2023 Medical Education Research Project of Henan Province, WJLX2023193; 2023 Henan Federation of Social Science research topic, SKL-2023-1102; Henan Province science and technology research project, 232102320062.

References

[1] M. M. Bendary, M. I. Abd El-Hamid, A. I. Abousaty, et al., Therapeutic Switching of Rafoxanide: a New Approach To Fighting Drug-Resistant Bacteria and Fungi. Microbiology spectrum, 2023, 12(2): 124-139. https://doi.org/10.1128/spectrum.02679-22

[2] D. Zhang, F. Yin, Q. Qin, & L. Qiao, Molecular responses during bacterial filamentation reveal inhibition methods of drug-resistant

bacteria. Proceedings of the National Academy of Sciences of the United States of America, 2023, 120(27): 12-27. https://doi.org/10.1073/pnas.2301170120

[3] M. Harris, Machine Learning and Artificial Intelligence for Pathogen Identification and Antibiotic Resistance Detection: Advancing Diagnostics for Urinary Tract Infections. BioMed, 2023, 3(2): 246-255. https://doi.org/10.3390/biomed3020022

[4] S. N. Helen, S. D. İlayda, A. E. Daniel, Direct microfluidic antibiotic resistance testing in urine with smartphone capture: significant variation in sample matrix interference between individual human urine samples. RSC advances, 2021, 11(60): 258-263. https://doi.org/10.1039/D1RA06867A

[5] A. Fergus, N. Tim, Antimicrobial use: importance of bacterial culture and susceptibility testing. In Practice, 2021, 43(9): 500-510. https://doi.org/10.1039/W1RA45168B

[6] T. Rui, L. Rui, T. Shiwei, et al., Machine learning in predicting antimicrobial resistance: a systematic review and meta-analysis. International journal of antimicrobial agents, 2022, 60(5-6): 625-645. https://doi.org/10.3290/antibiotics110568924

[7] J. K. In, M. Finlay, K. K. T., et al., Machine Learning for Antimicrobial Resistance Prediction: Current Practice, Limitations, and Clinical Perspective. Clinical microbiology reviews, 2022, 22(6). https://doi.org/12.1239/s41467-022-2593623.

[8] Y. Ren, T. Chakraborty, S. Doijad, et al., Prediction of antimicrobial resistance based on whole-genome sequencing and machine learning. Bioinformatics (Oxford, England), 2021, 38(2): 325-334. https://doi.org/10.1093/bioinformatics/btab681

[9] L. Murray, A. Boxall, J. Snape, et al., The selection and co-selection of antimicrobial resistance by non-antibiotic drugs and plant protection products. Access Microbiology, 2020, 2(7A). https://doi.org/10.2190/ s41467-022-259

[10] M. M. Magdalena, C. Carmen, R. Elena, et al., Urinary Tract Infections in Patients with Type 1 and Type 2 Diabetes: Etiology, Resistance to Antibacterial Chemicals and Virulence Features. Revista de Chimie, 2017, 68(3): 566-569. https://doi.org/10.19982/j.issn.1000-6621.20170161.

[11] A. Tabish, A. Sarfaraz, A. Muhammad, Artificial Intelligence for Antimicrobial Resistance Prediction: Challenges and Opportunities towards Practical Implementation. Antibiotics, 2023, 12(3). https://doi:10.13881/j.cnki.hljxmsy.2023.06.0095.

[12] A. G. Green, C. H. Yoon, M. L. Chen, et al., A convolutional neural network highlights mutations relevant to antimicrobial resistance in Mycobacterium tuberculosis. Nature communications, 2022, 13(1): 817-829. https://doi.org/10.1038/s41467-022-31236-0

[13] Özçelik Rıza, van Tilborg Derek, JiménezLuna José. Structure-based Drug discovery with Deep Learning. Chembiochem: a European journal of chemical biology, 2023, 12(8): 213-236. https://doi.org/10.19937/j.issn.1671-4199.2023.03.029.

[14] M. Jamshidi, O. Moztarzadeh, A. Jamshidi, et al., Future of Drug Discovery: The Synergy of Edge Computing, Internet of Medical Things, and Deep Learning. Future Internet, 2023, 15(4). https://doi.org/10.16431/j.cnki.1671-7236.2023.02.032.

[15] A. A. Keshavarzi, S. Milad, K. Heather, et al. Functional microRNA-Targeting Drug Discovery by Graph-Based Deep Learning. bioRxiv: the preprint server for biology, 2023. https://doi.org/10.16718/j.1009-7708.2023.05.020.

[16] K. D. Stergiou, G. M. Minopoulos, V. A. Memos, et al., A Machine Learning-Based Model for Epidemic Forecasting and Faster Drug Discovery. Applied Sciences, 2022, 12(21): 328-352. https://doi.org/10.3390/app122110766

[17] C. Hitesh, A. B. Amin, R. G. K., et al., Application of Artificial intelligence in Drug Discovery. Current pharmaceutical design, 2022. https://doi.org/10.16437/j.cnki.1007-5038.2022.03.016.

[18] K. Yanagisawa, M. Toratani, A. Asai, et al., Convolutional Neural Network Can Recognize Drug Resistance of Single Cancer Cells. International Journal of Molecular Sciences, 2020, 21(9): 316-328. https://doi.org/10.16718/j.1219-7908.2020.05.035.

[19] Z. N. Khan, J. Ahmad, Attention induced multi-head convolutional neural network for human activity recognition. Applied Soft Computing Journal, 2021, 11(6): 45-57. https://doi.org/10.1016/j.asoc.2021.107671

[20] R. Shubhankar, K. P. S. R, K. Vineet, A novel complex-valued convolutional neural network for medical image denoising. Biomedical Signal Processing and Control, 2021, 69. https://doi.org/10.25638/j.1009-8978.2021.05.032.

[21] T. Chao, L. Shuhua, D. Feng, et al., Image Reconstruction Based on Convolutional Neural Network for Electrical Resistance Tomography. IEEE Sensors Journal, 2019, 19(1):196-204. https://doi.org/10.1080/15257770.2019.2153251

[22] Y. Ülger, A. Delik, Artificial intelligence model with deep learning in nonalcoholic fatty liver disease diagnosis: genetic based artificial neural networks. Nucleosides, Nucleotides & Nucleic Acids, 2023, 42(5): 398-406. https://doi.org/10.1080/15257770.2022.2152046

[23] A. Tabish, A. Sarfaraz, A. Muhammad, Artificial Intelligence for Antimicrobial Resistance Prediction: Challenges and Opportunities towards Practical Implementation. Antibiotics, 2023, 12(3): 523-538. https://doi.org/10.1023/15145980.2023.269875

[24] S. Aikaterini, K. Christina, F. Georgios, et al. Using Machine Learning to Predict Antimicrobial Resistance—A Literature Review. Antibiotics, 2023, 12(3): 452-476. https://doi.org/10.16003/j.cnki.issn1672-5190.2023.03.006.

[25] A. A. Rabaan, S. Alhumaid, A. A. Mutair, et al., Application of Artificial Intelligence in Combating High Antimicrobial Resistance Rates. Antibiotics, 2022, 11(6): 784-798. https://doi.org/10.3390/antibiotics11060784

[26] L. Yuying, F. Tao, W. Yan, The role of bacterial signaling networks in antibiotics response and resistance regulation. Marine Life Science & Technology, 2022, 4(2):163-178. https://doi.org/10.16458/j.cnki.1007-0893.2022.06.002.

[27] B. Beatriz, A. C. Beatriz, M.S. P S, et al., Network biology and artificial intelligence drive the understanding of the multidrug resistance phenotype in cancer. Drug Resistance Updates, 2022, 60. https://doi.org/10.16437/j.cnki.1007-5038.2022.07.010.

[28] T. Hao, X. Chao, H. Xu, Electrical resistance tomography image reconstruction based on one-dimensional multi-branch convolutional neural network combined with attention mechanism. Flow Measurement and Instrumentation, 2022, 84. https://doi.org/10.19386/j.cnki.jxnyxb.2022.10.024.

[29] A. K. Bagherian, S. Sadaf, N. Mohammadjavad, et al., Point-of-Interest Preference Model Using an Attention Mechanism in a Convolutional Neural Network. Bioengineering (Basel, Switzerland), 2023, 10(4). https://doi.org/10.16485/j.issn.2095-7858.2023.01.007.

[30] W. Ziye, L. Shuo, Y. Ronghui, et al., ARG-SHINE: improve antibiotic resistance class prediction by integrating sequence homology, functional information and deep convolutional neural network. NAR genomics and bioinformatics, 2021, 3(3): 66-79. https://doi.org/10.16863/j.cnki.1003-6377.2021.04.011.

[31] Wei Tao. Research on the prediction method of drug resistance of nonsmall cell lung cancer based on deep learning. Guangxi University, Nanning, China. 2021. https://doi.org/10.27034/d.cnki.ggxiu.2020.002152.