



OPEN ACCESS

EDITED BY

Nguyen Quoc Khanh Le,
Taipei Medical University, Taiwan

REVIEWED BY

Adeel Malik,
Sangmyung University, Republic of Korea
Xin Liu,
Xuzhou Medical University, China

*CORRESPONDENCE

Muhammad Akmal Remli,
✉ akmal@umk.edu.my

RECEIVED 12 February 2025

ACCEPTED 23 May 2025

PUBLISHED 11 June 2025

CITATION

Wei X, Hu S, Tu J and Remli MA (2025) iKcr-DRC: prediction of lysine crotonylation sites in proteins based on a novel attention module and DenseNet.
Front. Genet. 16:1574832.
doi: 10.3389/fgene.2025.1574832

COPYRIGHT

© 2025 Wei, Hu, Tu and Remli. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

iKcr-DRC: prediction of lysine crotonylation sites in proteins based on a novel attention module and DenseNet

Xin Wei^{1,2,3}, Siqin Hu⁴, Jian Tu² and Muhammad Akmal Remli^{1,3*}

¹Institute for Artificial Intelligence and Big Data, Universiti Malaysia Kelantan, Kota Bharu, Kelantan, Malaysia, ²Business School, Jiangxi Institute of Fashion Technology, Nanchang, China, ³Faculty of Data Science and Computing, Universiti Malaysia Kelantan, Kota Bharu, Kelantan, Malaysia, ⁴School of Mega Data, Jiangxi Institute of Fashion Technology, Nanchang, China

Introduction: Lysine crotonylation (Kcr) is a recently identified post-translational modification that predominantly occurs on lysine residues and plays a crucial role in regulating gene expression, cellular metabolism, and various biological processes. Increasing evidence has linked Kcr to the pathogenesis of major diseases such as cancer, highlighting the importance of accurately identifying Kcr sites for understanding disease mechanisms and normal cellular function.

Methods: In this study, we present a novel deep learning-based computational model, named iKcr-DRC, for the accurate prediction of lysine crotonylation sites. The model leverages a densely connected convolutional network (DenseNet) as its backbone to effectively capture high-level local features from protein sequences. Additionally, we introduce an enhanced channel attention mechanism with a short-circuit connection design, endowing the network with residual properties and improved feature refinement capabilities.

Results: The experimental results show that the iKcr-DRC model achieves 90.30%, 78.35%, 84.33% and 69.15% for sensitivity, specificity, accuracy, and Matthew's correlation coefficients, respectively. These results indicate a significant improvement over existing state-of-the-art Kcr prediction tools.

Discussion: The proposed iKcr-DRC model provides an effective and innovative approach for predicting lysine crotonylation sites. It holds great potential for advancing applications in bioinformatics and enhancing the understanding of protein post-translational modifications. An online prediction tool based on the iKcr-DRC model is freely accessible at: <http://www.lzzzlab.top/ikcr/>.

KEYWORDS

protein post-translational modification, lysine crotonylation site, deep learning, DenseNet, channel attention mechanism

1 Introduction

Protein post-translational modifications (PTMs) (Hu et al., 2023) are widespread physiological phenomena in organisms and play an important role in the functional regulation of cells (Kouzarides, 2007; Verdin and Ott, 2015). Currently, modern technology has discovered more than 600 different types of PTMs (Meng et al., 2022), which include a variety of modifications such as phosphorylation, glycolysis, acetylation and succination (Macek et al., 2019). These modifications can change protein structure,

function, and interaction, thus producing important regulatory effects on biological processes within the cell (Fu et al., 2018). Protein lysine crotonylation (Kcr) (Tan et al., 2011) is an important PTM modality that occurs on lysine residues. Catalyzed by acyltransferases, acetyl groups are added to residues of lysine, thereby changing protein chemical structure and function. Kcr plays an important role in several cells processes, including key processes such as gene transcription and cellular metabolism (Patel et al., 2024). It is involved in the regulation of these cells processes by modulating protein interactions, activity, and stability (Hou et al., 2021). Medical research has shown that Kcr plays an important regulatory role in major diseases such as tumors and cancer (Gao et al., 2023; Jiang et al., 2021). Therefore, accurate prediction of the Kcr site is critical for understanding an organism's normal function.

In the past, In order to explore the importance of Kcr sites in biological processes, researchers have developed a number of biological experimental methods (Yu et al., 2020) to identify Kcr sites. However, these methods suffer from high costs. Therefore, researchers urgently need to develop more convenient and efficient computational methods to replace the traditional biomedical experimental methods. Computational methods enable the prediction of Kcr sites, while biological experiments provide validation and contextual support for these predictions. In recent years, researchers have developed a series of compelling Kcr site models. These models can accurately predict Kcr sites of proteins, thus helping researchers to further explore the functions and regulatory mechanisms of protein modifications in biology. In 2016, researchers created a small dataset of Kcr sites and applied a discrete hidden Markov model (DHMM) to predict them (Qiu et al., 2017). Subsequently, many researchers (Ju and He, 2017; Liu et al., 2020; Malebary et al., 2019; Qiu et al., 2017; Qiu et al., 2018; Wang et al., 2020) have worked based on this dataset. In 2017, Qiu et al. (2017) developed a model called Position-weight. The model encodes the Kcr site sequence using position-weighted amino acid composition (PWAA) (Ismail et al., 2016) and predicts it by support vector machine (SVM). The results of the study show that the model exhibits good performance in prediction. In the same year, Ju and He (2017) developed a model called CKSAAP-CrotSite, which utilizes k-spaced amino acid pairs (CKSAAP) and SVM to encode and predict Kcr sites. In 2019, the iKcr-PseEns model (Qiu et al., 2018) and the iCrotoK-PseAAC model (Malebary et al., 2019) used Random Forest and Artificial Neural Network (ANN) to predict Kcr sites, respectively. These two models employ different algorithms and feature coding methods to further improve the accuracy and performance of the Kcr site prediction models. In 2020, Liu et al. (2020) developed a model called LightGBM-CroSite. The model encodes Kcr site sequences with multiple complicated coding methods and uses the LightGBM algorithm for prediction. Experimental results showed that the model demonstrated the best prediction results. These models performed well in predicting the Kcr site, especially the iCrotoK-PseAAC and LightGBM-CroSite models, which had a prediction accuracy of 99%. However, machine learning predictive models usually perform well on small datasets, but they are relatively less innovative. In addition, these machine learning models are highly dependent on complex feature encoding methods to improve predictive performance, yet the complexity of feature coding is a rather difficult task. All these limitations are drawbacks of current machine learning models.

Deep learning models (Sun et al., 2022; Wang T. et al., 2023; Wang et al., 2022; Zhang L. et al., 2021) show great potential in a big data-driven context. In 2020, Lv et al. (2021a) created a balanced benchmark dataset of Kcr sites, while they developed a predictive model called Deep-Kcr by using a convolutional neural network. In 2021, Qiao et al. (2022) based on this dataset, they encoded the Kcr site sequences by the BERT model (Le et al., 2021) and utilized the BiLSTM network (Sharma et al., 2022) to capture the global features of the Kcr site sequences, so as to extract richer feature information for predicting the Kcr site. Experimental results show that this model achieves good results. In the same year, Khanal et al. (2022) introduced capsule network (Chen Z. et al., 2023) to replace the traditional convolutional neural network, because the traditional convolutional neural network may lose the spatial information when compressing the feature information, and the introduction of capsule network further improved the experimental results. However, the main network structures adopted by these prediction models are too basic and lack innovation, and thus there is still a wide scope of exploration in the field of deep learning methods.

We propose a deep learning model called iKcr-DRC, which employs DenseNet (Wei et al., 2022) as the core network structure to extract advanced local feature information. Traditional convolutional neural networks pass feature information from layer by layer, whereas DenseNet networks use a densely connected structure that allows each layer to be directly connected to all previous layers, thus effectively utilizing the feature map information of the previous layers. This dense connection structure allows our model to avoid network degradation problems and extract richer feature information. In addition, we have made innovative improvements to the channel attention mechanism (Chen W. et al., 2023; Meng et al., 2023) by introducing the design of short-circuit connections, which makes the channel attention mechanism internally equipped with residual structures. The improved channel attention mechanism can more accurately compute the channel weights of the output feature maps of the DenseNet network, thus improving the performance and accuracy of the model. With these improvements, our iKcr-DRC model shows better performance in the Kcr site prediction task. This study provides new ideas and methods for the application of deep learning in bioinformatics.

2 Materials and methods

2.1 Benchmark dataset

In this study, we used the benchmark dataset created by Lv et al. (2021b). The protein sequences in this dataset were extracted from the UniProt database (UniProt Consortium, 2011), and to avoid redundancy, they applied the CD-HIT tool (Huang et al., 2010) to remove duplicate protein sequences with more than 30% similarity. In the end, they obtained 9,964 samples of real Kcr site and 9,964 samples of spurious Kcr site. Based on experimental confirmation, the appropriate length for positive and negative samples is 29, and they are both centered at the K site. In positive samples, K indicates crotonylation, while in negative samples, K indicates no crotonylation modification. Finally, they

TABLE 1 The details of the benchmark dataset.

Original dataset	Positive	Negative
Training Dataset	6,975	6,975
Testing Dataset	2,989	2,989
Total	9,964	9,964

randomly split the positive and negative samples into a training set and an independent test set in a ratio of 7:3. The final training dataset contained 6,975 Kcr site samples and 6,975 non-Kcr site samples, while the independent test dataset contained 2,989 Kcr site samples and 2,989 non-Kcr site samples. To ensure the uniqueness of each sequence, different classes of protein sequences are mutually independent. The details of the benchmark dataset are shown in Table 1.

2.2 Feature extraction methods

There are some shortcomings in the current methods of protein feature coding. First, some feature coding methods encode protein sequences as fixed-length vectors or matrices, but these methods can be improved in capturing contextual information between proteins. Second, traditional coding methods rely on manual design, which leads to complex and difficult feature construction. In summary, these traditional methods have some limitations in protein feature coding.

Word embedding model (Ren et al., 2022) is a feature coding model based on contextual semantic information and shallow neural network training. In the field of bioinformatics, word embedding models are widely applied. Word embedding coding of protein sequences involves the following steps: first, the protein sequence is divided into K-mers of fixed length using a sliding window. Next, these k-mers are transformed to binary vectors using One-hot coding (Abbas et al., 2022). All binary vectors are trained using a two-layer neural network to generate the weight parameter matrix. Finally, this weight parameter matrix is multiplied with a binary vector to obtain a dense vector for representing K-mer words. By training the word embedding model, we can obtain a dense vector representation of each K-mer word, and it possesses contextual semantic relations. We set the sliding window size to 1, the fixed length to 1-mers, and use an 80-dimensional word vector to represent each 1-mer. Therefore, the length of each Krc site sequence is 29, which can be expressed as a feature matrix of size 80×29 .

2.3 Model construction

In this study, we propose a deep learning prediction model called iKcr-DRC, which is capable of automatically coding protein sequences and extracting high-level feature information to improve the prediction accuracy. First, we apply a word embedding technique to each Krc site sequence and convert it into a feature encoding matrix of size 80×29 . Then, we directly input this feature encoding matrix into the DenseNet network with Dense Connectivity to fully explore the deep features of the sequences. Subsequently, we introduce the residual channel attention mechanism on the high-level feature maps extracted by

DenseNet, which further enhances the representation of the feature maps by emphasizing the importance of the information of each channel. Finally, we input the feature maps evaluated by the attention mechanism into the fully connected neural network for prediction. Figure 1 illustrates the specific details of this network framework.

2.3.1 DenseNet

In this study, we make structural improvements based on the standard DenseNet framework. The traditional DenseNet usually contains an initial convolutional layer, multiple Dense blocks, and several Transition layers, where low-level features are first extracted from the initial convolutional layer, and then input into the Dense blocks and Transition layers to obtain richer high-level representations. In our improved scheme, the initial convolutional layer is first eliminated, and the feature coding matrix is directly input into the Dense blocks to fully utilize the information of the original data. At the same time, we also add an additional batch normalization layer (Berrar and Dubitzky, 2021) between each Dense block and Transformation layer, which improves the robustness and stability of the model while reducing its complexity.

2.3.1.1 Dense block

Dense block (Yang et al., 2020) is a key component of DenseNet network and it plays an important role. The Dense block consists of multiple densely connected convolutional layers, where the output of each convolutional layer is connected to the outputs of all previous convolutional layers to produce a densely connected network structure. The Dense block equation is shown below.

$$x_l = F_l([x_0, x_1, \dots, x_{l-1}]), l = 1, 2, \dots, L. \quad (1)$$

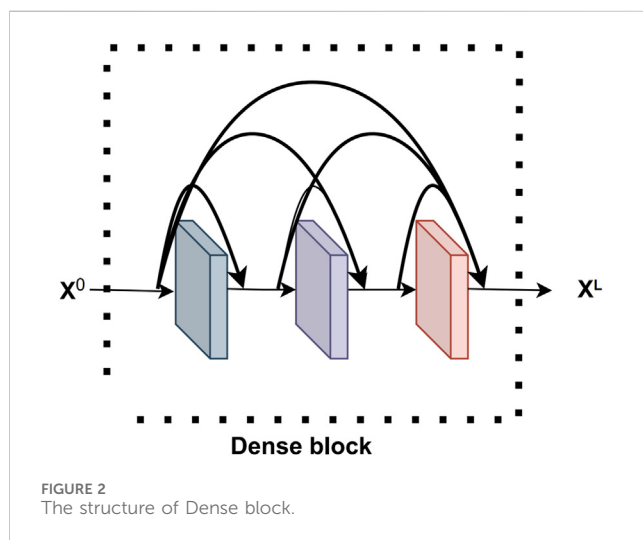
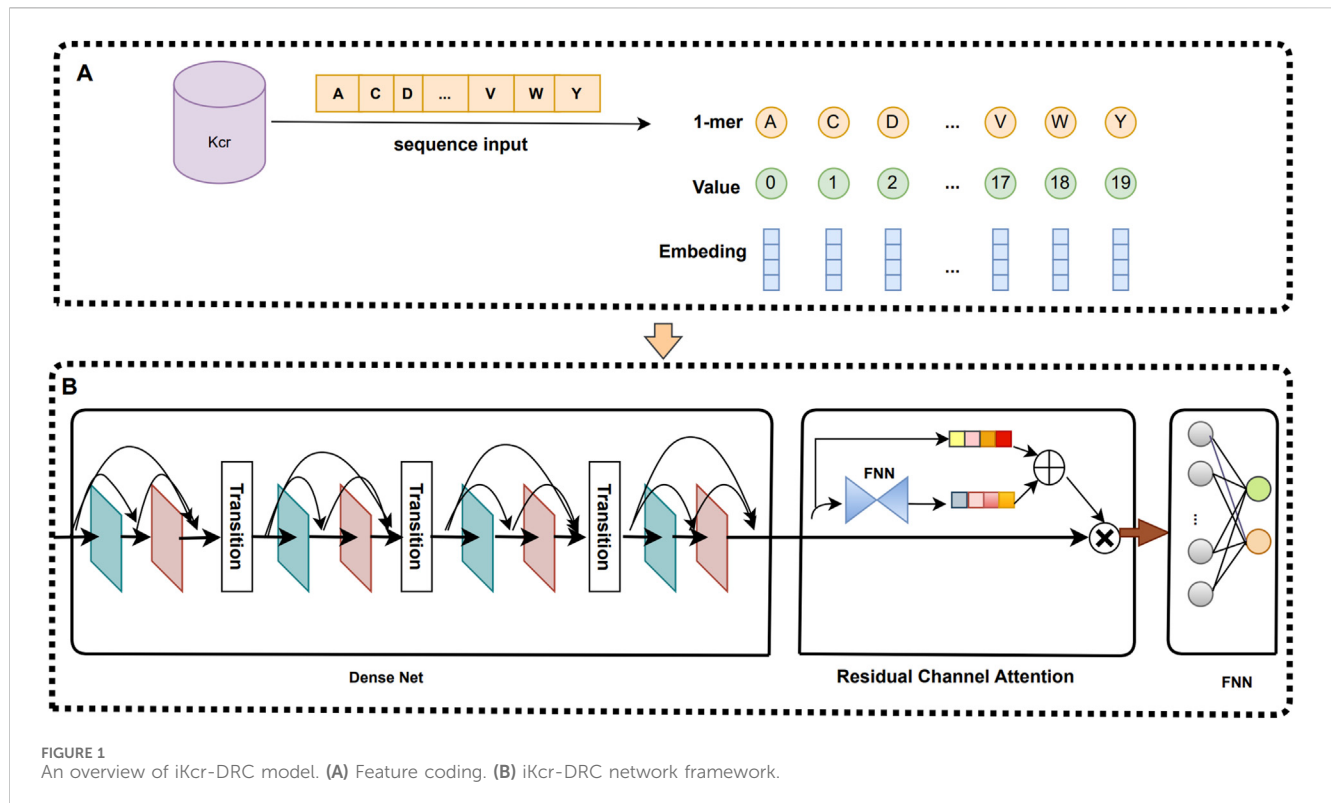
$$F_l(x) = \text{Conv}(\text{ReLU}(\text{BN}(x))) \quad (2)$$

where x_l denotes the concatenation operation in the feature channel dimension and F is a composite function consisting of the batch normalization (BN), ReLU activation function and convolution (Conv) operation.

This densely connected structure can effectively integrate shallow features with deeper features, thus realizing the reuse of feature information. The Dense block structure improves the overall performance of the model in terms of gradient flow, computational efficiency, and model generalization ability, enabling iKcr-DRC to perform Krc site prediction more accurately. The Dense block structure is shown in Figure 2.

2.3.1.2 Dense block

The Transition layer (Jia et al., 2023a) in the DenseNet network is mainly designed to reduce the size of the feature map and thus control the complexity of the model. The Transition layer is located between adjacent Dense blocks and consists of a convolutional layer (1×1) and an average pooling layer (2×2). The parameters of the Dense block's output feature maps are huge, which can lead to parameter explosion and slower training. Therefore, the Transition layer uses a convolutional layer (1×1) to reduce channel number of the feature map and an average pooling layer (2×2) to compress the spatial size of the feature map. This reduces the parameters of the feature map and reduces the complexity of the model, thus improving the generalizability of the model.



We add a batch normalization layer to the Transition layer. The batch normalization layer can normalize the input data to optimize and accelerate the model training and improve the model performance. The Transition layer equation is shown below.

$$x_{trans} = T(x_{block}) = \text{AvgPool}(\text{Conv}_{1 \times 1}(\text{BN}([x_0, \dots, x_L]))) \quad (3)$$

2.3.2 Residual channel attention mechanisms

In convolutional neural networks, the weights of each channel of the feature map are fixed, and there is no adaptive learning based on the importance of the features. However, the importance of feature information is not the same for different channels. By using the

channel attention mechanism, we can adaptively weight the channel feature information according to its importance, thus enhancing the attention of the model on important feature information. This improves the representation of features, which in turn improves the performance and generalization of the model. The structure of the channel attention mechanism is shown in Figure 3.

We assume that in the channel attention mechanism, useless feature information is already filtered when the feature map is dimensionally compressed using the Pooling layer. In a way, the feature vectors obtained by pooling layer compression already has the potential role of channel attention weights. Based on this assumption, we improve the channel attention mechanism by introducing a short-circuit connection structure to improve its performance. The improved channel attention mechanism is called residual channel attention mechanism. The residual channel attention mechanism equation is shown below.

$$M_c(F) = \sigma(\text{FNN}(\text{AvgPool}(F)) + \text{FNN}(\text{MaxPool}(F))) \quad (4)$$

$$M_r(F) = \sigma(\text{AvgPool}(F) + \text{MaxPool}(F)) \quad (5)$$

$$M_{rc}(F) = M_c(F) + M_r(F) \quad (6)$$

$$F = F_{scale}(F, M_{rc}(F)) = M_{rc}(F) \cdot F \quad (7)$$

where the pooling here is global max pooling and global average pooling. $M_{rc}(F)$ denotes the residual weight value. F_{scale} denotes each channel specific value F multiplied by the weight.

In this residual channel attention mechanism, we add the feature vectors compressed by the pooling layer with the weight vectors obtained from the network training. In this way, the generated channel weight vectors can more accurately represent the importance of each channel in the feature map, thus improving representation of key feature information by the model. The purpose

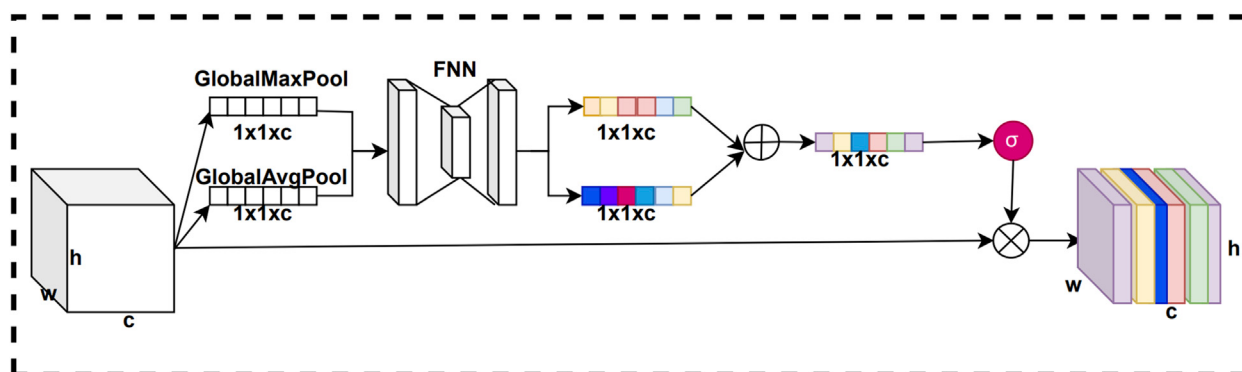


FIGURE 3
Structure of the channel attention mechanism.

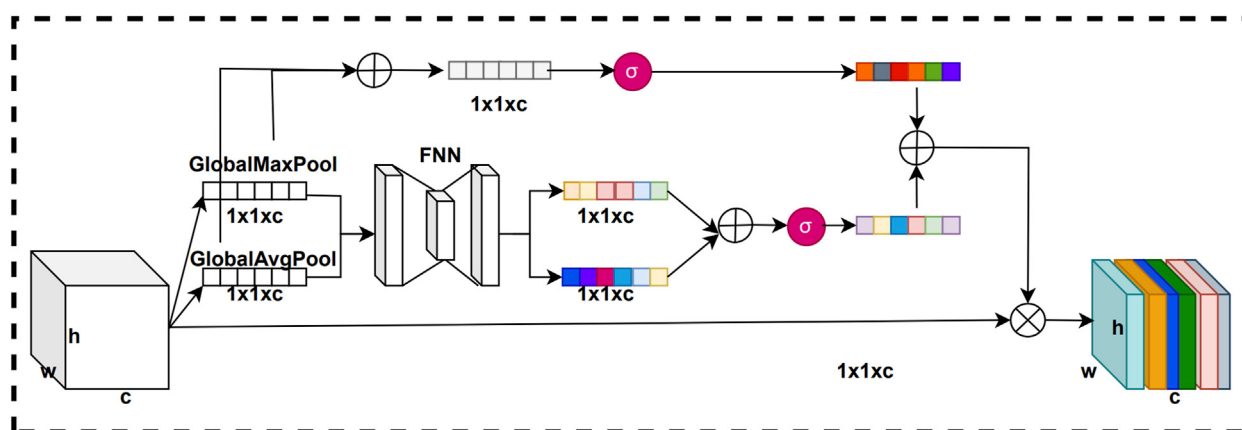


FIGURE 4
Structure of the residual channel attention mechanism.

of this improvement is to make the network more attentive to important channel features and thus improve the performance of the model. The structure of the residual channel attention mechanism is shown in Figure 4.

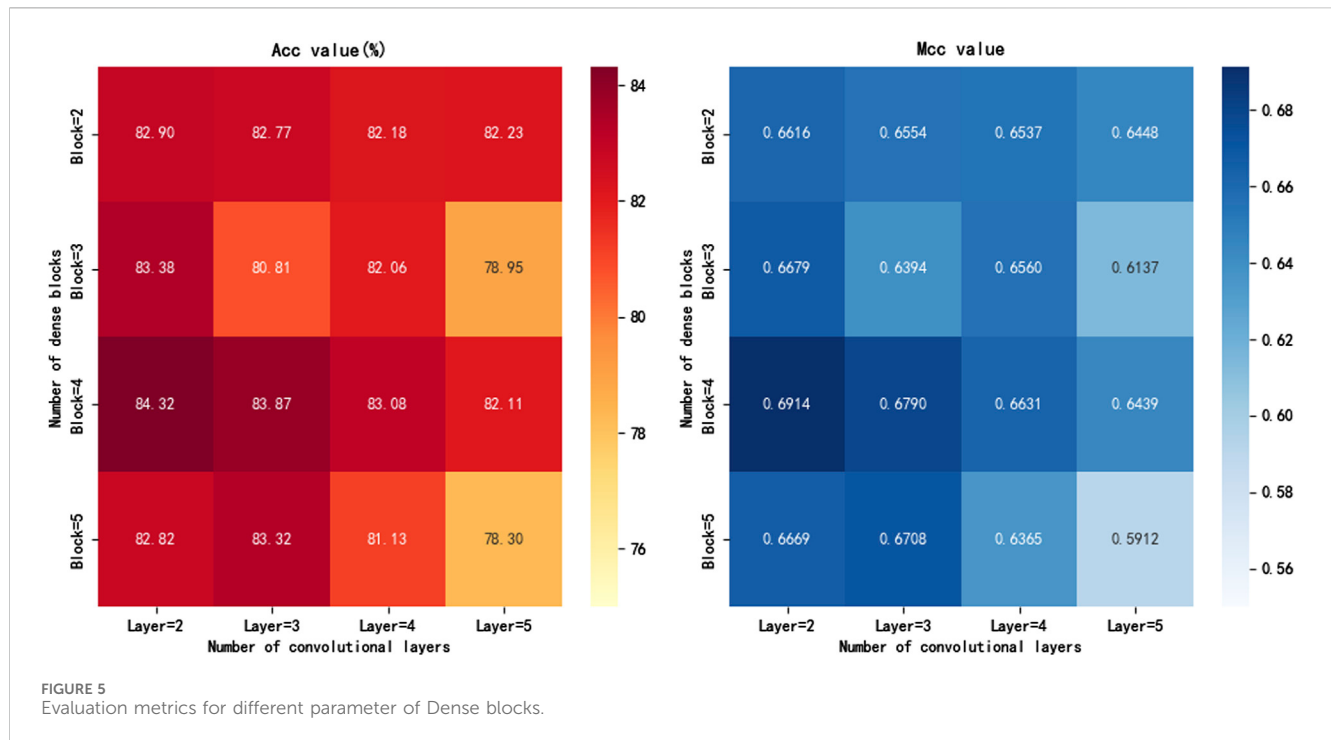
The original channel attention mechanism first uses global max pooling and global average pooling compresses on the feature map to obtain pooled vectors of two channel dimension sizes. Next, we input these two pooled vectors into a fully connected neural network for training to get two trained vectors. Then, we sum these two training vectors to get a summation vector. Finally, the summation vector is normalized using the sigmoid activation function to obtain the channel attention weights. To improve the channel attention mechanism, we introduce a short-circuit connection structure. First, we add the two pooling vectors to obtain a summation pooling vector. Next, we process the summation pooling vector using a sigmoid activation function to obtain a residual weight vector. Finally, the residual weight vector is summed with the channel attention weights to obtain the final residual channel attention weights.

2.3.3 Fully connected neural network

We use DenseNet and residual channel attention mechanism as advanced feature extraction methods. With these methods, we can obtain a set of advanced feature representations. Next, the extracted advanced features are flattened into vectors and fed into a Fully connected neural network. Finally, we employ the softmax function (Duhan et al., 2022) to predict classification results.

2.4 Performance evaluation

To evaluate the predictive performance of the model more comprehensively, we used four common scientific evaluation metrics for the assessment. These evaluation metrics include sensitivity (Sn), specificity (Sp), accuracy (Acc) and Matthew's correlation coefficient (MCC) (Jia et al., 2023b). By using these evaluation metrics, we can understand the experimental effect of the model in detail, which helps us to further improve and optimize the model. The evaluation metrics are as in Equation 8.



$$\begin{cases}
 Sp = \frac{TN}{TN + FP} \\
 Sn = \frac{TP}{TP + FN} \\
 Acc = \frac{TP + TN}{TP + TN + FP + FN} \\
 MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}}
 \end{cases}
 \quad (8)$$

Among them, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are the four important indicators in the confusion matrix (Niu et al., 2021). These metrics represent the model's accuracy in predicting positive and negative samples, respectively, and they are the basis for computing the evaluation metrics. To characterize the model performance more accurately, we also used ROC curves and AUC metrics (Jia et al., 2022) for evaluation. Higher values of all these evaluation metrics indicate better model performance.

3 Results and discussion

3.1 Comparing parameter combinations of dense blocks

In this study, DenseNet serves as the core network structure of the iKcr-DRC model, and we conduct an in-depth exploration of the parameter combinations of the Dense block to effectively avoid overfitting while ensuring model performance. Specifically, we compare the model performance by adjusting the number of Dense blocks and

the number of convolutional layers inside them. According to the experimental results in Figure 5, when the number of Dense blocks is set to four and the number of convolutional layers in each Dense block is set to 2, the Acc and MCC metrics of the model reach their maximum values. Based on this finding, we finally chose this parameter configuration to build the DenseNet network to achieve the best prediction results while balancing the complexity and performance.

3.2 Comparison of different coding methods

In this study, we compare several matrix feature coding methods because our model network structure can only accept feature matrices as input. These matrix feature coding methods include Word embedding, One-hot and AAindex (Wang X. et al., 2023). In addition, we discuss the application of Word embedding coding in three cases: 1-mer, 2-mer and 3-mer. We also compare the current popular protein large language model ESM-2. We input these matrix feature codes into the network framework and then compare the performance of the models. According to the experimental results in Figure 6, Word embedding coding has the highest value of all evaluation metrics when comparing the other three coding methods. The One-hot coding method has problems of sparsity and information loss, and the AAindex coding method mainly focuses on the amino acid composition of proteins while ignoring other key structural and functional information. The ESM-2 coding method generates embedding vectors of protein sequences based on a large-scale pre-trained language model, which is theoretically able to capture the deep semantic and structural information in the sequences. However, limited by computational resources, the ESM-2 model

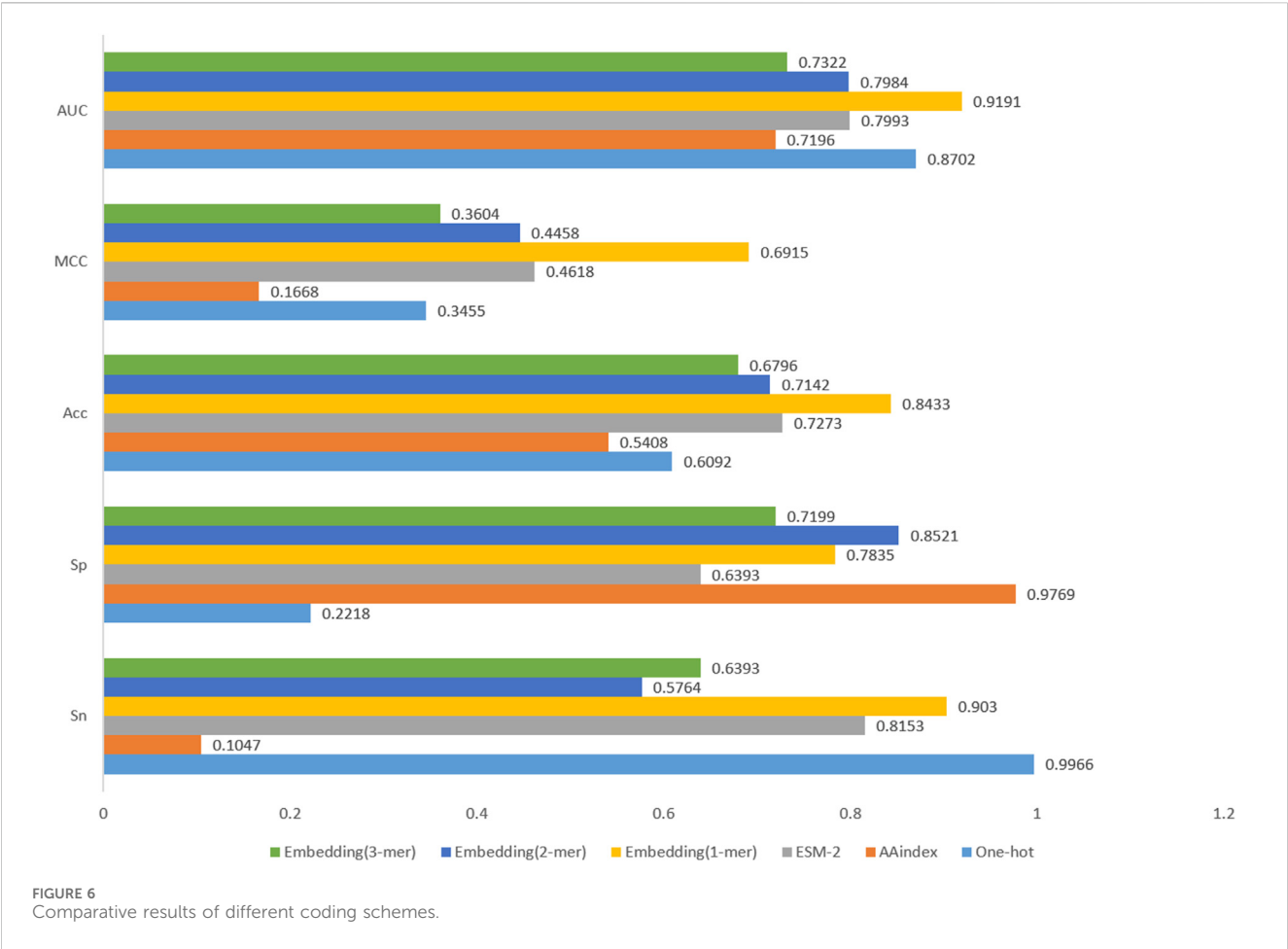


TABLE 2 Comparison of different attention mechanisms.

Model frameworks	Sn	Sp	Acc	MCC	AUC
SENet	0.8855	0.7785	0.8320	0.6679	0.9169
CAM	0.9203	0.7236	0.8220	0.6568	0.9126
ECANet	0.7460	0.8909	0.8185	0.6437	0.9106
CBAM	0.9260	0.7363	0.8312	0.6746	0.9181
DANet	0.7708	0.8728	0.8218	0.6470	0.9137
RCAM	0.9030	0.7835	0.8433	0.6915	0.9191

The highest score for each metric is highlighted by bold.

with the smallest parameters was chosen for encoding in this study, which led to its relatively poor performance in the experiments. In contrast, word embedding coding is a method of encoding word vectors based on contextual semantic information by network training, which is more efficient than manual feature coding. Meanwhile, we compare the experimental results of Word embedding coding in three cases (1-mer, 2-mer and 3-mer), and the results show that simple 1-mer obtains the best results. Therefore, we identified 1-mer Word embedding coding as the preferred scheme for feature coding.

3.3 Comparison of different attention mechanisms

In this study, we researched and analyzed the six attentional mechanisms. These six mechanisms include SENet (Li et al., 2023), channel attention mechanism (CAM), ECANet (Gong et al., 2023), CBAM (Jia et al., 2023c), DANet (Bai et al., 2023) and residual channel attention mechanism (RCAM). We combined these attention mechanisms with Dense Net and tested them. According to the results of the test experiments in Table 2, we

TABLE 3 Comparison of different network architecture models.

Model frameworks	Sn	Sp	Acc	MCC	AUC
DenseNet	0.9488	0.6306	0.7897	0.6112	0.9021
ResNet + RCAM	0.8752	0.6814	0.7783	0.5674	0.8503
InceptionResNet + RCAM	0.9973	0.1063	0.5518	0.2283	0.8146
EfficientNet + RCAM	0.7142	0.6219	0.6681	0.3376	0.7148
BiGRU	0.8159	0.8166	0.8163	0.6326	0.8884
Transformer	0.8591	0.7969	0.8280	0.6573	0.9077
ViT	0.8390	0.7039	0.7714	0.5480	0.8438
DenseNet + RCAM	0.9030	0.7835	0.8433	0.6915	0.9191

The highest score for each metric is highlighted by bold.

can see that the RCAM shows better experimental results compared to the CAM. This suggests that by introducing a short-circuit connection structure into the CAM, the model can highlight more important channel feature information. Compared to simple SENet and CBAM, the RCAM is slightly better in terms of performance. The two latest attention mechanisms, ECANet and DANet, exhibit relatively poor performance, this is because complex attention mechanisms are usually only fully utilized on large-scale datasets. Synthesizing the experimental results, we finally chose the Residual Channel Attention Mechanism.

3.4 Comparison of different network frameworks

We researched the performance of multiple deep learning network frameworks in the task of Kcr site prediction. Specifically, we researched convolutional neural network models such as ResNet (Ge et al., 2022), DenseNet, InceptionResNet (Zhang X. et al., 2021), and EfficientNet (Dehghan Rouzi et al., 2023), and further experimented with two temporal network frameworks, BiGRU (Zhang et al., 2023) and Transformer (Zhou et al., 2022). In addition, we have tried the Vision Transformer (ViT) (Khedr et al., 2023) model, which is very compelling in the field of computer vision. According to the experimental results shown in Table 3, compared to using only the DenseNet network, the performance is significantly improved by adding the RCAM. InceptionResNet and EfficientNet are two newer deep neural network models that employ a series of complex structural designs to improve model performance. However, the simple DenseNet network structure shows better performance compared to these two complex convolutional neural networks. This indicates that the use of more complex network structures does not necessarily improve performance on small datasets. Similarly, we tried the BiGRU and Transformer models. These models showed good performance in the experiments, but they were slightly less effective relative to the combination of DenseNet and RCAM. ViT is a large model and requires more data to fully utilize its superior performance. However, in small data samples, the ViT network model cannot be utilized to its full potential. This indicates that by using DenseNet and the RCAM to extract features, we can obtain more advanced feature information, which makes it easier to predict the Kcr site. Based on the

experimental results, we choose the combination of DenseNet and RCAM to build the final network framework.

3.5 Comparing different combinations of attention mechanisms and DenseNet networks

In this study, we delve into three different approaches for combining residual channel attention mechanisms with DenseNet networks. The goal of these combined approaches is to fully utilize the advantages of the residual channel attention mechanism and the DenseNet network to further enhance model performance and feature representation. These three different combinations are shown in Figure 7. The first combination approach is called Dense + RCAM + Net, which adds the residual channel attention mechanism between the Dense blocks and the Transient layer. The second combination approach is called Den + RCAM + seNet, which introduces the residual channel attention mechanism between the Convolutional layers of each Dense block. The third combination is called DenseNet + RCAM, which adds the residual channel attention mechanism after the final output of DenseNet.

After researching and experimenting with the three combination methods, we arrived at the experimental results shown in Figure 8. Surprisingly, we found that the simplest DenseNet + RCAM combination obtained the best results. We analyzed the effects of these two complex combinations approaches, but found that they led to a complication of the original DenseNet network and triggered an overfitting problem. In contrast, DenseNet + RCAM has less changes in the model structure, which may help to retain DenseNet's original feature extraction capability and information mobility. This could also explain the better performance of DenseNet + RCAM, because it better balances the complexity and the generalization ability of the model.

3.6 Performance of iKcr-DRC on training dataset

Cross-validation is a very important means to validate the performance of a model. To validate the performance of the iKcr-

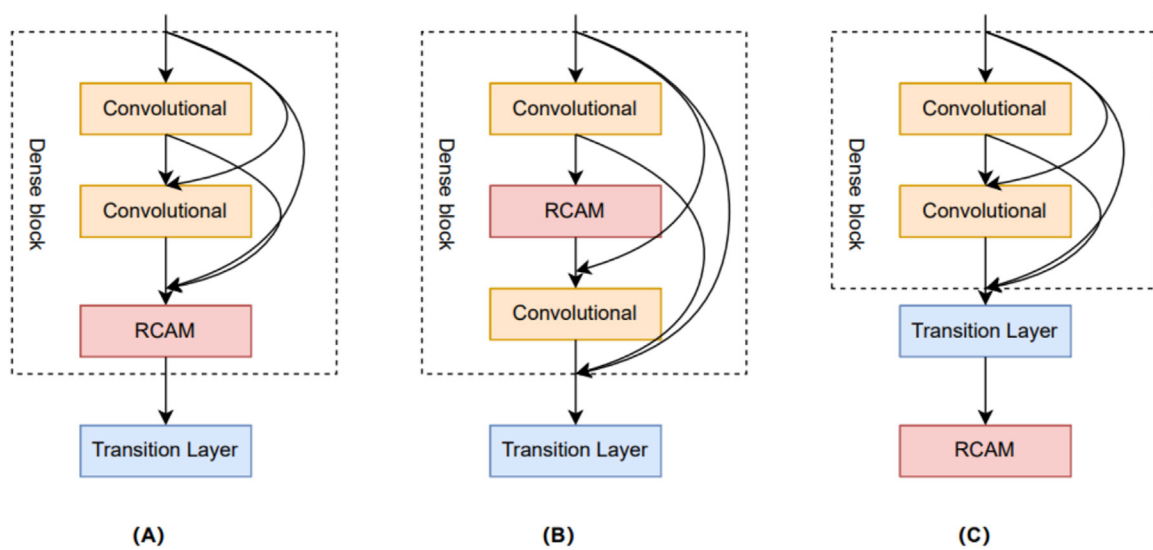


FIGURE 7
Three different network combination structures. (A) Dense + RCAM + Net. (B) Den + RCAM + seNet. (C) DenseNet + RCAM.

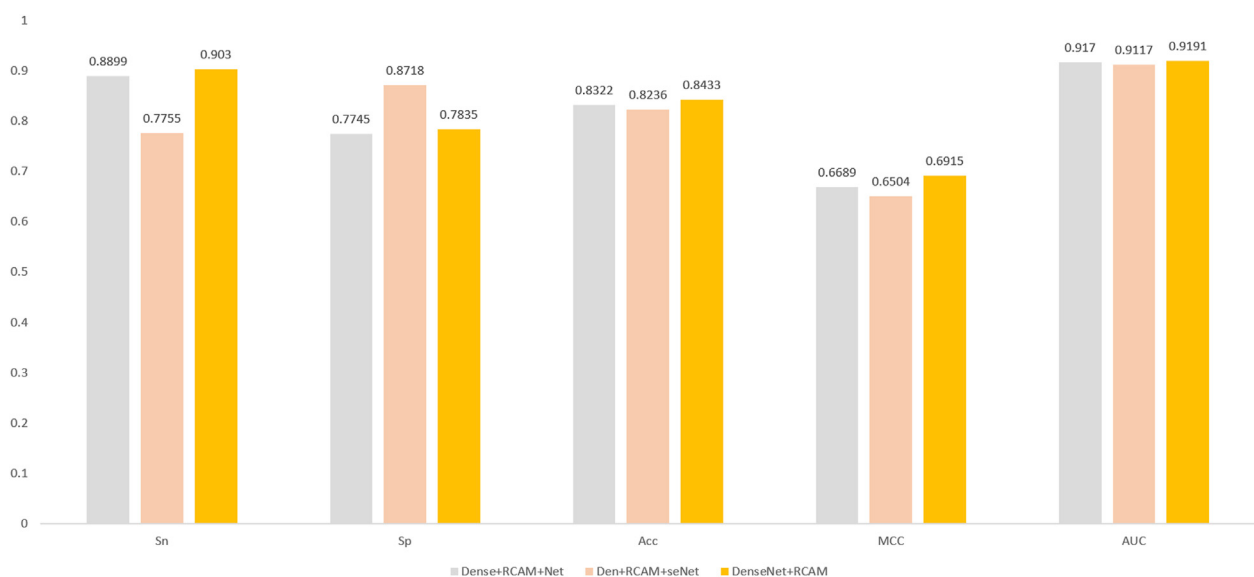
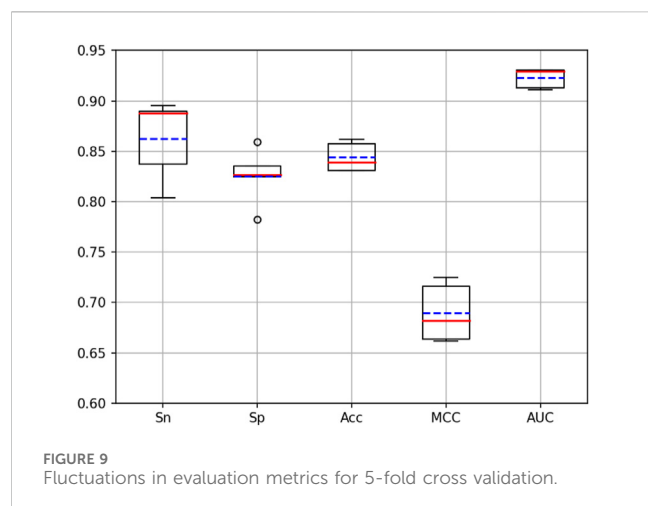


FIGURE 8
Experimental results for three different network combination structures.

DRC model in the task of predicting Kcr site, we performed a 5-fold cross-validation on the training dataset. According to the results shown in Figure 9, the values of other evaluation metrics show relatively stable fluctuations when performing the 5-fold cross-validation, except for two outliers in the Sp metric. This proves that the iKcr-DRC model has more stable and superior performance. In addition, we also compared the average results of the iKcr-DRC model with the DeepCap-Kcr model for 5-fold cross-validation, and the comparison results are displayed in Figure 10. The iKcr-DRC model improved Sn, Acc and MCC by 4.97%, 1.01% and 1.56%, respectively. This indicates that the iKcr-DRC model has good robustness in the task of predicting Kcr site.

3.7 Comparison of iKcr-DRC with existing predictors

In recent years, many excellent computational methods have emerged in the field of Kcr sites. To comprehensively evaluate the performance of the iKcr-DRC model, we performed independent tests on benchmark datasets and compared it with existing models. In this way, we can better understand the performance of the iKcr-DRC model on the benchmark dataset. We show the experimental results in Table 4 and found that the iKcr-DRC model significantly outperforms the



existing models on four important evaluation metrics (Sn, Acc, MCC and AUC). Specifically, in independent tests, the iKcr-DRC model improved Sn, Acc, MCC, and AUC by 7.9%, 2.03%, 3.55%, and 0.91%, respectively, compared to the most advanced DeepCap-Kcr model. This indicates that the iKcr-DRC model has superior performance and can more accurately and reliably predict Kcr sites. Although the iKcr-DRC model performs very well, it may have a lower Sp metric value compared to other models. This is because we are embedding coding (1-mer), which works better for positive samples, but may slightly reduce the Sp

metric value while maintaining high accuracy. Nevertheless, the iKcr-DRC model still has excellent performance and shows higher accuracy in positive sample prediction. And in the field of bioinformatics, accurate prediction of positive samples is a very important task, so our model and coding method are more meaningful in the field of bioinformatics.

Analyzing other predictors, Deep-Kcr is a predictor developed based on traditional convolutional neural networks, however, traditional convolutional neural networks are relatively low in performance. Although BERT-Kcr utilizes the BERT model and BiLSTM network to capture global features of Kcr site sequences, it is deficient in extracting local feature information. In addition, DeepCap-Kcr employs a capsule network for prediction, which compensates for some of the shortcomings of CNNs, but the performance improvement is limited. In contrast, DenseNet extracts more advanced local information by iteratively utilizing previous features, thus becoming the key to the superior performance of this model.

3.8 Comparison of iKcr-DRC with existing predictors in other datasets

Our model employs an adaptive coding method that is capable of accomplishing coding based on the contextual semantic information of the sequence. The core goal of this experiment was to validate the migratory and cross-task adaptation of the iKcr-DRC model, not for predicting other

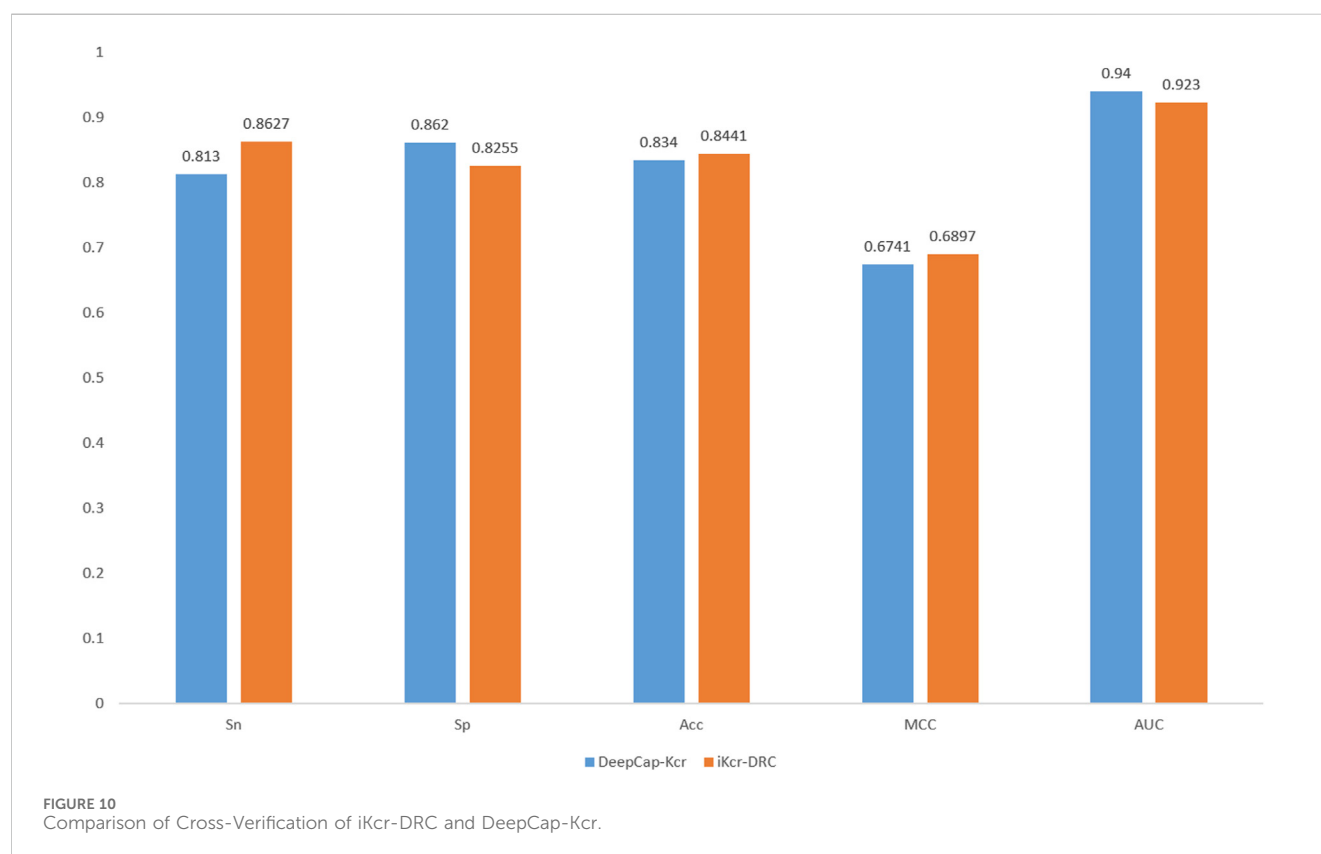


TABLE 4 Compared with other models on the same independent datasets.

Predictor model	Sn	Sp	Acc	MCC	AUC
Deep-Kcr	0.630	0.871	0.751	0.516	0.859
BERT-Kcr	0.801	0.838	0.820	0.640	0.905
DeepCap-Kcr	0.824	0.836	0.823	0.656	0.910
iKcr-DRC	0.9030	0.7835	0.8433	0.6915	0.9191

The highest score for each metric is highlighted by bold.

TABLE 5 Comparison of the serine/threonine (S/T) phosphorylation site dataset with other models.

Predictor model	Sn	Sp	Acc	MCC	AUC
DeepPSP	0.7665	0.8378	0.8021	0.606	0.876
Bert-ST	0.8007	0.7460	0.7984	0.600	0.889
DeepIPs	0.7961	0.8350	0.8063	0.632	0.894
iKcr-DRC	0.8090	0.8341	0.8215	0.643	0.910

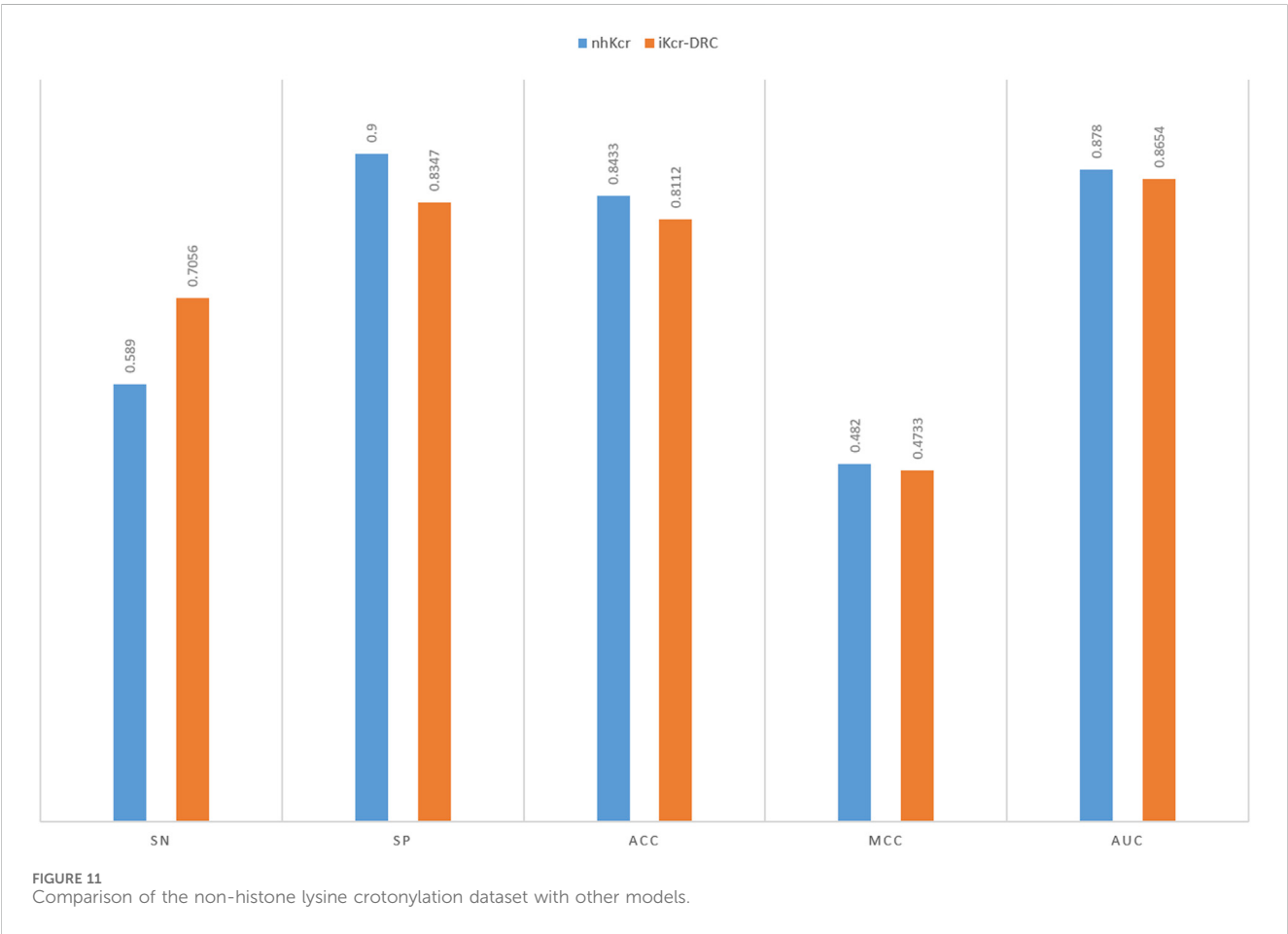
The highest score for each metric is highlighted by bold.

PTM tasks. In this experiment, we did not make any adjustments to the hyperparameters of the iKcr-DRC model, but trained and tested it directly on the training set and independent test set of

other PTMs. Despite the differences in the biological mechanisms of different types of PTMs, we hypothesize that certain common structural or functional features may be embedded in the sequence contexts, thus giving the model some potential for cross-task recognition. In addition, to further validate the performance of the iKcr-DRC model, we tested the iKcr-DRC model again on the serine/threonine (S/T) phosphorylation site dataset of [Lv et al. \(2021b\)](#). The benchmark dataset contains 5,387 positive and 5,387 negative samples, and they randomly divide the positive and negative samples into a training set and an independent test set in the ratio of 8:2. The results of the independent tests are shown in [Table 5](#), where our model significantly outperforms the other three prediction models in the Sn, Acc, MCC and AUC metrics. Our model improves the most important MCC metrics by 1.1%–4.3%. It means that the iKcr-DRC model shows excellent robustness in the prediction task of protein post-translational modification and is expected to become the most representative protein post-translational modification prediction model.

3.9 Prediction of iKcr-DRC in the crotonylation of non-histone lysine

Compared to the widely studied lysine crotonylation of histones, non-histone Kcr modifications have shown broader



biological significance in the field of functional proteomics. Kcr modifications of non-histone proteins have more diverse regulatory functions in organelles and cellular processes, and its sequence structure is more complex and exhibits a high degree of heterogeneity. As a result, this type of dataset faces significant challenges in the feature learning process, which puts the generalization ability of the predictive model to a great test. Accurate identification of Kcr sites on non-histone proteins not only helps to reveal their potential functions beyond transcriptional regulation, but also sets a higher standard for constructing computational models with good generalization and robustness. Specifically, we tested the iKcr-DRC model on the non-histone lysine crotonylation dataset from Chen Z. et al. (2021). The benchmark dataset contains 15,605 positive samples and 75,111 negative samples, and they randomly divided the positive and negative samples into a training set and an independent test set in the ratio of 8:2. Note that this benchmark dataset is unbalanced. The results of the independent tests are shown in Figure 11, where our model slightly decreases by less than 1% on the most important MCC metric. However, on the Sn metric, our model improves by 11%. In the unbalanced dataset, a higher Sn metric indicates that the model exhibits higher accuracy in the prediction of positive samples. This result re-validates the effectiveness of our model and coding method in positive sample prediction and further demonstrates their importance in the field of bioinformatics. With this extended study, we can more comprehensively assess the performance and potential of the iKcr-DRC model, which will provide an important reference and guidance for further development in the field of non-histone lysine crotonylation research.

4 Conclusion

This study is devoted to the development of a new Kcr sites prediction model named iKcr-DRC. We employ the DenseNet network as the core network framework to extract advanced local feature information of Kcr sites. At the same time, we have improved the Channel Attention Mechanism so that it can better highlight important channel feature information. The iKcr-DRC model significantly outperforms other existing models in all evaluation metrics, which proves its excellent expressiveness and performance. We are committed to combining deep learning methods with Kcr site research to drive progress in the Kcr field. The iKcr-DRC model shows potential not only in Kcr site prediction, but also has a value of application in predicting serine/threonine (S/T) phosphorylation sites. This means that the iKcr-DRC model provides researchers with a convenient tool to help advance the field of post-translational modification of proteins.

Although iKcr-DRC performs well on balanced data, it still has shortcomings on unbalanced datasets and requires further research and improvement. Unfortunately, the interpretability of the iKcr-DRC model has not been systematically investigated, which will be one of the core research elements in our follow-up work. As we continue to study the Kcr site, we expect that more advanced deep learning methods will bring more possibilities for Kcr site prediction.

Data availability statement

The data presented in the study are deposited in the <https://github.com/weixin7112/iKcr-DRC>.

Author contributions

XW: Data curation, Methodology, Writing - original draft. SH: Validation, Writing - original draft. JT: Investigation, Writing - original draft. MR: Supervision, Writing - review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was partially supported by the National Natural Science Foundation of China (Nos. 61761023, 62162032, and 31760315), the Natural Science Foundation of Jiangxi Province, China (Nos. 20202BABL202004 and 20202BAB202007), the Scientific Research Plan of the Department of Education of Jiangxi Province, China (GJJ2402711 and GJJ2402712). These funders had no role in the study design, data collection and analysis, decision to publish or preparation of manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2025.1574832/full#supplementary-material>

References

- Abbas, Z., Tayara, H., Chong, K. T., Siddeeq, S. M., and Jung, J. (2022). ENet-6mA: identification of 6mA modification sites in plant genomes using ElasticNet and neural networks. *Int. J. Mol. Sci.* 23 (15), 15266. doi:10.3390/ijms232315266
- Bai, Y., Li, J., Shi, L., Jiang, Q., Yan, B., and Wang, Z. (2023). DME-DeepLabV3+: a lightweight model for diabetic macular edema extraction based on DeepLabV3+ architecture. *Front. Med. (Lausanne)* 10, 1150295. doi:10.3389/fmed.2023.1150295
- Berrari, D., and Dubitzky, W. (2021). Deep learning in bioinformatics and biomedicine. *Briefings Bioinforma.* 22 (2), 1513–1514. doi:10.1093/bib/bbab087
- Chen, W., Zhang, R., Zhang, Y., Bao, F., Lv, H., Li, L., et al. (2023). Pact-Net: parallel CNNs and Transformers for medical image segmentation. *Comput. Methods Programs Biomed.* 242, 107782. doi:10.1016/j.cmpb.2023.107782
- Chen, Z., Zhang, L., Sun, J., Meng, R., Yin, S., and Zhao, Q. (2023). DCAMCP: a deep learning model based on capsule network and attention mechanism for molecular carcinogenicity prediction. *J. Cell Mol. Med.* 27 (20), 3117–3126. doi:10.1111/jcmm.17889
- Dehghan Rouzi, M., Moshiri, B., Khoshnevisan, M., Akhaee, M. A., Jaryani, F., Salehi Nasab, S., et al. (2023). Breast cancer detection with an ensemble of deep learning networks using a consensus-adaptive weighting method. *J. Imaging* 9 (11), 247. doi:10.3390/jimaging9110247
- Duhan, N., Norton, J. M., and Kaundal, R. (2022). deepNEC: a novel alignment-free tool for the identification and classification of nitrogen biochemical network-related enzymes using deep learning. *Briefings Bioinforma.* 23 (3), bbac071. doi:10.1093/bib/bbac071
- Fu, J., Wu, M., and Liu, X. (2018). Proteomic approaches beyond expression profiling and PTM analysis. *Anal. Bioanal. Chem.* 410 (17), 4051–4060. doi:10.1007/s00216-018-1021-y
- Gao, H., Sun, J., Wang, Y., Lu, Y., Liu, L., Zhao, Q., et al. (2023). Predicting metabolite–disease associations based on auto-encoder and non-negative matrix factorization. *Briefings Bioinforma.* 24 (5), bbad259. doi:10.1093/bib/bbad259
- Ge, F., Zhang, Y., Xu, J., Muhammad, A., Song, J., and Yu, D. J. (2022). Prediction of disease-associated nsSNPs by integrating multi-scale ResNet models with deep feature fusion. *Briefings Bioinforma.* 23 (1), bbab530. doi:10.1093/bib/bbab530
- Gong, R., He, S., Tian, T., Chen, J., Hao, Y., and Qiao, C. (2023). FRCNN-AA-CIF: an automatic detection model of colon polyps based on attention awareness and context information fusion. *Comput. Biol. Med.* 158, 106787. doi:10.1016/j.compbiomed.2023.106787
- Hou, J. Y., Zhou, L., Li, J. L., Wang, D. P., and Cao, J. M. (2021). Emerging roles of non-histone protein crotonylation in biomedicine. *Cell Biosci.* 11 (1), 101. doi:10.1186/s13578-021-00616-2
- Hu, H., Feng, Z., Lin, H., Cheng, J., Lyu, J., Zhang, Y., et al. (2023). Gene function and cell surface protein association analysis based on single-cell multiomics data. *Comput. Biol. Med.* 157, 106733. doi:10.1016/j.compbiomed.2023.106733
- Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010). CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26 (5), 680–682. doi:10.1093/bioinformatics/btq003
- Ismail, H. D., Newman, R. H., and Kc, D. B. (2016). RF-Hydroxysite: a random forest based predictor for hydroxylation sites. *Mol. Biosyst.* 12 (8), 2427–2435. doi:10.1039/c6mb00179c
- Jia, J., Lei, R., Qin, L., Wu, G., and Wei, X. (2023a). iEnhancer-DCSV: predicting enhancers and their strength based on DenseNet and improved convolutional block attention module. *Front. Genet.* 14, 1132018. doi:10.3389/fgene.2023.1132018
- Jia, J., Sun, M., Wu, G., and Qiu, W. (2023b). DeepDN_iGlu: prediction of lysine glutarylation sites based on attention residual learning method and DenseNet. *Math. Biosci. Eng.* 20 (2), 2815–2830. doi:10.3934/mbe.2023132
- Jia, J., Wei, Z., and Cao, X. (2023c). EMDL-ac4C: identifying N4-acetylcytidine based on ensemble two-branch residual connection DenseNet and attention. *Front. Genet.* 14, 1232038. doi:10.3389/fgene.2023.1232038
- Jia, J., Wu, G., Li, M., and Qiu, W. (2022). pSuc-EDBAM: predicting lysine succinylation sites in proteins based on ensemble dense blocks and an attention module. *BMC Bioinforma.* 23 (1), 450. doi:10.1186/s12859-022-05001-5
- Jiang, G., Li, C., Lu, M., Lu, K., and Li, H. (2021). Protein lysine crotonylation: past, present, perspective. *Cell Death Dis.* 12 (7), 703. doi:10.1038/s41419-021-03987-z
- Ju, Z., and He, J. J. (2017). Prediction of lysine crotonylation sites by incorporating the composition of k-spaced amino acid pairs into Chou's general PseAAC. *J. Mol. Graph Model* 77, 200–204. doi:10.1016/j.jmgm.2017.08.020
- Khanal, J., Tayara, H., Zou, Q., and To Chong, K. (2022). DeepCap-Kcr: accurate identification and investigation of protein lysine crotonylation sites based on capsule network. *Briefings Bioinforma.* 23 (1), bbab492. doi:10.1093/bib/bbab492
- Khedr, O. S., Wahed, M. E., Al-Attar, A. R., and Abdel-Rehim, E. A. (2023). The classification of the bladder cancer based on Vision Transformers (ViT). *Sci. Rep.* 13 (1), 20639. doi:10.1038/s41598-023-47992-y
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128 (4), 693–705. doi:10.1016/j.cell.2007.02.005
- Le, N. Q. K., Ho, Q.-T., Nguyen, T.-T.-D., and Ou, Y. Y. (2021). A transformer architecture based on BERT and 2D convolutional neural network to identify DNA enhancers from sequence information. *Briefings Bioinforma.* 22 (5), bbab005. doi:10.1093/bib/bbab005
- Li, H., Shen, C., Wang, G., Sun, Q., Yu, K., Li, Z., et al. (2023). BloodNet: an attention-based deep network for accurate, efficient, and costless bloodstain time since deposition inference. *Briefings Bioinforma.* 24 (1), bbac557. doi:10.1093/bib/bbac557
- Liu, Y., Yu, Z., Chen, C., Han, Y., and Yu, B. (2020). Prediction of protein crotonylation sites through LightGBM classifier based on SMOTE and elastic net. *Anal. Biochem.* 609, 113903. doi:10.1016/j.ab.2020.113903
- Lv, H., Dao, F. Y., Guan, Z. X., Yang, H., Li, Y. W., and Lin, H. (2021b). Deep-Kcr: accurate detection of lysine crotonylation sites using deep learning method. *Brief. Bioinform* 22 (4), bbab255. doi:10.1093/bib/bbaa255
- Lv, H., Dao, F.-Y., Zulfiqar, H., and Lin, H. (2021a). DeepIPs: comprehensive assessment and computational identification of phosphorylation sites of SARS-CoV-2 infection using a deep learning-based approach. *Briefings Bioinforma.* 22 (6), bbab244. doi:10.1093/bib/bbab244
- Macek, B., Forchhammer, K., Hardouin, J., Weber-Ban, E., Grangeasse, C., and Mijakovic, I. (2019). Protein post-translational modifications in bacteria. *Nat. Rev. Microbiol.* 17 (11), 651–664. doi:10.1038/s41579-019-0243-0
- Malebary, S. J., Rehman, M. S. U., and Khan, Y. D. (2019). iCrotoK-PseAAC: identify lysine crotonylation sites by blending position relative statistical features according to the Chou's 5-step rule. *PLoS One* 14 (11), e0223993. doi:10.1371/journal.pone.0223993
- Meng, L., Chan, W. S., Huang, L., Liu, L., Chen, X., Zhang, W., et al. (2022). Mini-review: recent advances in post-translational modification site prediction based on deep learning. *Comput. Struct. Biotechnol. J.* 20, 3522–3532. doi:10.1016/j.csbj.2022.06.045
- Meng, R., Yin, S., Sun, J., Hu, H., and Zhao, Q. (2023). scAAGA: single cell data analysis framework using asymmetric autoencoder with gene attention. *Comput. Biol. Med.* 165, 107414. doi:10.1016/j.compbiomed.2023.107414
- Niu, K., Luo, X., Zhang, S., Teng, Z., Zhang, T., and Zhao, Y. (2021). iEnhancer-EBLSTM: identifying enhancers and strengths by ensembles of bidirectional long short-term memory. *Front. Genet.* 12, 665498. doi:10.3389/fgene.2021.665498
- Patel, A. B., He, Y., and Radhakrishnan, I. (2024). Histone acetylation and deacetylation - mechanistic insights from structural biology. *Gene* 890, 147798. doi:10.1016/j.gene.2023.147798
- Qiao, Y., Zhu, X., and Gong, H. (2022). BERT-Kcr: prediction of lysine crotonylation sites by a transfer learning method with pre-trained BERT models. *Bioinformatics* 38 (3), 648–654. doi:10.1093/bioinformatics/btab712
- Qiu, W.-R., Sun, B.-Q., Tang, H., Huang, J., and Lin, H. (2017). Identify and analysis crotonylation sites in histone by using support vector machines. *Artif. Intell. Med.* 83, 75–81. doi:10.1016/j.artmed.2017.02.007
- Qiu, W. R., Sun, B. Q., Xiao, X., Xu, Z. C., Jia, J. H., and Chou, K. C. (2018). iKcr-PseEn: identify lysine crotonylation sites in histone proteins with pseudo components and ensemble classifier. *Genomics* 110 (5), 239–246. doi:10.1016/j.ygeno.2017.10.008
- Ren, R., Yin, C., and S. S. T. Y. (2022). kmer2vec: a novel method for comparing DNA sequences by word2vec embedding. *J. Comput. Biol.* 29 (9), 1001–1021. doi:10.1089/cmb.2021.0536
- Sharma, R., Shrivastava, S., Kumar Singh, S., Saxena, S., and Kumar Singh, R. (2022). Deep-AFPpred: identifying novel antifungal peptides using pretrained embeddings from seq2vec with 1DCNN-BiLSTM. *Briefings Bioinforma.* 23 (1), bbab422. doi:10.1093/bib/bbab422
- Sun, F., Sun, J., and Zhao, Q. (2022). A deep learning method for predicting metabolite–disease associations via graph neural network. *Briefings Bioinforma.* 23 (4), bbac266. doi:10.1093/bib/bbac266
- Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., et al. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146 (6), 1016–1028. doi:10.1016/j.cell.2011.08.008
- UniProt Consortium (2011). Ongoing and future developments at the universal protein resource. *Nucleic Acids Res.* 39, D214–D219. doi:10.1093/nar/gkq1020
- Verdin, E., and Ott, M. (2015). 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat. Rev. Mol. Cell Biol.* 16 (4), 258–264. doi:10.1038/nrm3931
- Wang, R., Wang, Z., Wang, H., Pang, Y., and Lee, T. Y. (2020). Characterization and identification of lysine crotonylation sites based on machine learning method on both plant and mammalian. *Sci. Rep.* 10 (1), 20447. doi:10.1038/s41598-020-77173-0

- Wang, T., Sun, J., and Zhao, Q. (2023). Investigating cardiotoxicity related with hERG channel blockers using molecular fingerprints and graph attention mechanism. *Comput. Biol. Med.* 153, 106464. doi:10.1016/j.combiomed.2022.106464
- Wang, W., Zhang, L., Sun, J., Zhao, Q., and Shuai, J. (2022). Predicting the potential human lncRNA-miRNA interactions based on graph convolution network with conditional random field. *Briefings Bioinforma.* 23 (6), bbac463. doi:10.1093/bib/bbac463
- Wang, X., Ding, Z., Wang, R., and Lin, X. (2023). Deepr-Glu: combination of convolutional neural network and Bi-LSTM models using ProtBert and handcrafted features to identify lysine glutarylation sites. *Brief. Bioinform* 24 (2), bbac631. doi:10.1093/bib/bbac631
- Wei, P. J., Pang, Z. Z., Jiang, L. J., Tan, D. Y., Su, Y. S., and Zheng, C. H. (2022). Promoter prediction in nannochloropsis based on densely connected convolutional neural networks. *Methods* 204, 38–46. doi:10.1016/j.ymeth.2022.03.017
- Yang, T., Wu, T., Li, L., and Zhu, C. (2020). SUD-GAN: deep convolution generative adversarial network combined with short connection and dense block for retinal vessel segmentation. *J. Digit. Imaging* 33 (4), 946–957. doi:10.1007/s10278-020-00339-9
- Yu, H., Bu, C., Liu, Y., Gong, T., Liu, X., Liu, S., et al. (2020). Global crotonylome reveals CDYL-regulated RPA1 crotonylation in homologous recombination-mediated DNA repair. *Sci. Adv.* 6 (11), eaay4697. doi:10.1126/sciadv.aay4697
- Zhang, L., Yang, P., Feng, H., Zhao, Q., and Liu, H. (2021). Using network distance analysis to predict lncRNA-miRNA interactions. *Interdiscip. Sci.* 13 (3), 535–545. doi:10.1007/s12539-021-00458-z
- Zhang, M., Gao, H., Liao, X., Ning, B., Gu, H., and Yu, B. (2023). DBGRU-SE: predicting drug-drug interactions based on double BiGRU and squeeze-and-excitation attention mechanism. *Briefings Bioinforma.* 24 (4), bbad184. doi:10.1093/bib/bbad184
- Zhang, X., Li, Y., Tao, Y., Wang, Y., Xu, C., and Lu, Y. (2021). A novel method based on infrared spectroscopic inception-resnet networks for the detection of the major fish allergen parvalbumin. *Food Chem.* 337, 127986. doi:10.1016/j.foodchem.2020.127986
- Zhou, Y., Wang, X., Yao, L., and Zhu, M. (2022). LDAformer: predicting lncRNA-disease associations based on topological feature extraction and Transformer encoder. *Briefings Bioinforma.* 23 (6), bbac370. doi:10.1093/bib/bbac370