Chinese Journal of Electronics
vol. 33, no. 1, pp. 256–263, January 2024
https://doi.org/10.23919/cje.2022.00.361

RESEARCH ARTICLE

An Encoding-Decoding Framework Based on CNN for circRNA-RBP Binding Sites Prediction

Yajing GUO[1], Xiujuan LEI[1], and Yi PAN[2,3]

1. *School of Computer Science, Shaanxi Normal University, Xi'an 710119, China*
2. *Faculty of Computer Science and Control Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China*
3. *Department of Computer Science, Georgia State University, Atlanta, GA 30302, USA*

Corresponding author: Xiujuan LEI, Email: xjlei@snnu.edu.cn
Manuscript Received November 1, 2022; Accepted April 17, 2023

Abstract — Predicting RNA binding protein (RBP) binding sites on circular RNAs (circRNAs) is a fundamental step to understand their interaction mechanism. Numerous computational methods are developed to solve this problem, but they cannot fully learn the features. Therefore, we propose circ-CNNED, a convolutional neural network (CNN)-based encoding and decoding framework. We first adopt two encoding methods to obtain two original matrices. We preprocess them using CNN before fusion. To capture the feature dependencies, we utilize temporal convolutional network (TCN) and CNN to construct encoding and decoding blocks, respectively. Then we introduce global expectation pooling to learn latent information and enhance the robustness of circ-CNNED. We perform circ-CNNED across 37 datasets to evaluate its effect. The comparison and ablation experiments demonstrate that our method is superior. In addition, motif enrichment analysis on four datasets helps us to explore the reason for performance improvement of circ-CNNED.
Keywords — Circular RNAs (circRNAs), RNA binding proteins, Convolutional neural network, Temporal convolutional network, Encoder-decoder network.
Citation — Yajing GUO, Xiujuan LEI, Yi PAN, "An Encoding-Decoding Framework Based on CNN for circRNA-RBP Binding Sites Prediction," *Chinese Journal of Electronics*, vol. 33, no. 1, pp. 256–263, 2024. doi: 10.23919/cje.2022.00.361.

I. Introduction

Circular RNAs (circRNAs) are characterized by covalent closure [1]. They are strongly associated with biological processes [2] and diseases [3], [4] especially cancers [5]. They can also serve as microRNA (miRNA) sponges [5] and biomarkers [6]. circRNAs can regulate cancers by binding RNA binding proteins (RBPs) [7]. The binding events between them play important roles in some cellular processes [8]–[10]. Therefore, analyzing the interaction mechanism of them is pivotal and the binding sites prediction is helpful for this.

With the advent of many biological experimental methods [11]–[13], various databases related to circRNAs have been built. For example, circBase [14] collects circRNA information of multiple species. CircInteractome [15] stores the RBP binding sites information on circRNAs. There are other databases such as Circ2Traits [16], circRNADb [17], and so on [18]. Benefit from these databases, numerous computational methods are proposed.

Compared with traditional biological methods, computational methods based on deep learning are obviously more advantageous [19] and have been used in bioinformatics [20]–[22]. There are many computational methods to predict RBP binding sites on circRNAs [23]. For example, CRIP [24] constructs a hybrid neural network with a stacked encoding scheme based on codon. CRPBsites [25] comes up with a long short-term memory (LSTM)-based architecture. HCRNet [26] adopts deep temporal convolutional network (TCN) architecture [27]. Besides, a web tool CirRBP [28], is recently proposed. However, most methods still need improvement since the insufficient learning of feature information.

Therefore, we propose an encoding and decoding framework named circ-CNNED whose structure is shown as Figure 1. It applies convolutional neural network (CNN)

Associate Editor: Prof. Ran SU, Tianjin University.

and TCN to predict whether a circRNA sequence is a RBP binding site. To capture original features of circRNA sequences, we employ one-hot method to get a raw matrix. Considering the characteristics of circRNAs, we apply loop-3-mer to get another raw matrix to incorporate context information. Innovatively, the two matrices are preprocessed with CNN for concatenation. Next, the concatenated matrix is processed by a TCN encoder block and a CNN decoder block. Then, we utilize a global expectation pooling layer. Finally, a fully connected layer (FC)-softmax structure is used to classify. On 37 datasets, we compare circ-CNNED with four latest methods to demonstrate its effectiveness and the experimental results indicate that circ-CNNED has superior performance.

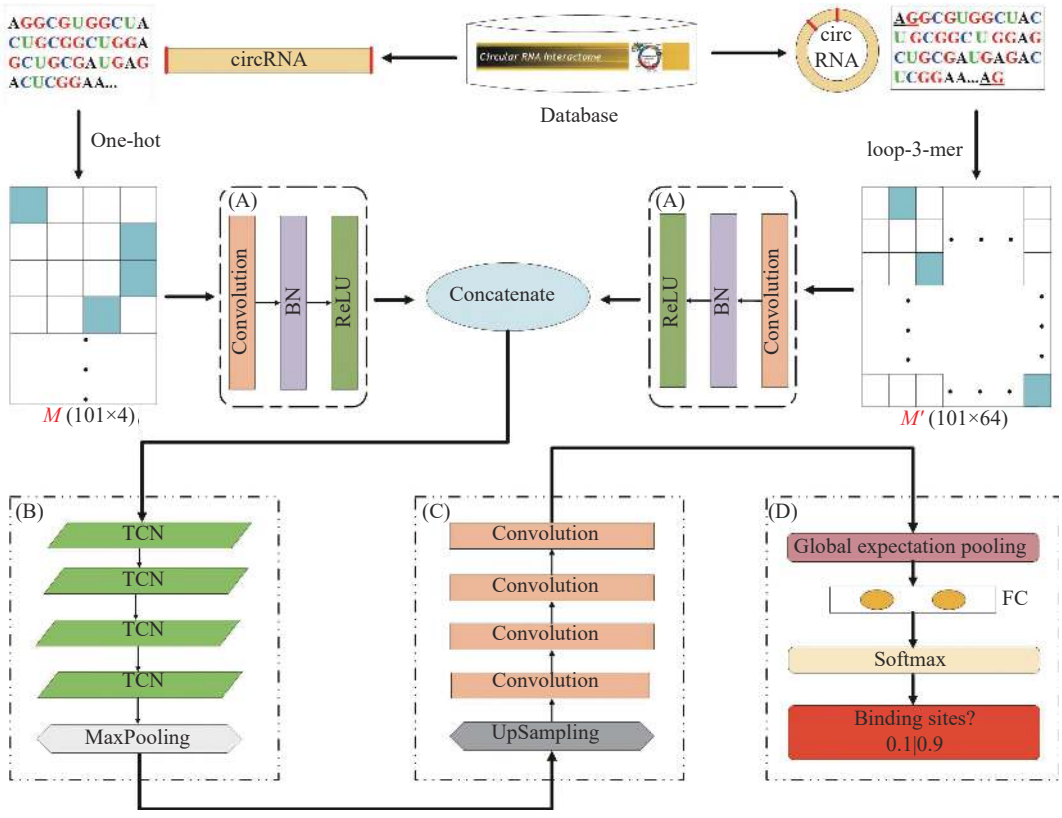

Figure 1 The structure of circ-CNNED. (A) Data preprocessing; (B) TCN encoder block; (C) CNN decoder block; (D) Classification block.

II. Materials and Methods

1. Datasets

To illustrate the efficiency of circ-CNNED, a dataset consisting of 37 sub-datasets corresponding to 37 RBPs is built [24]. The circRNA sequences are downloaded from CircInteractome database. From each CLIP-seq peak, we extend 50 nt (nucleotide) upstream and downstream respectively to get 101 nt segments as positive samples. For negative samples, we extract 101 nt segments from remaining fragments of circRNAs randomly. Then through CD-HIT [29] whose threshold is 0.8, the redundant sequences are eliminated. The number of positive and negative samples within each sub-dataset are equal for data balance. Finally, we have a dataset associated with 37 RBPs.

2. Sequence encoding

We use one-hot and loop-3-mer to get two raw matrices. One-hot method is able to not only maintain sequence raw information to the greatest extent, but also obtain a low-dimensional matrix. Given a circRNA sequence $S = \{s_1, s_2, s_3, \ldots, s_L\}$ whose length is L. s_i is the ith nucleotide, $i = \{1, 2, 3, \ldots, L\}$. Through one-hot, the circRNA sequence S is encoded as a matrix $\boldsymbol{M} = (m_{i,j})_{L\times 4}$ whose elements are represented as

$$m_{i,j} = \begin{cases} 1, & \text{if } s_i = j\text{th element } of\ \{A, U, C, G\} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where i is the row and j is the column.

Considering the circular structure of circRNAs, we adopt loop-3-mer to get another matric whose encoding rules are shown in Figure 2.

First, we complete the first two nucleotides of the circRNA sequence at the end of the sequence. Then we can get a new circRNA sequence $S' = \{s_1, s_2, s_3, \ldots, s_L, s_1, s_2\}$ whose length is $L + 2$. Considering the structure of amino acids, we then use 3-mer to encode S'. A window whose length is 3 slides along S' with stride=1 and S' is separated into L segments. One-hot method then transfers S' into a $L \times 64$ matrix $\boldsymbol{M}'$.

3. CNN for data preprocessing

Through sequence encoding, we get two matrices $\boldsymbol{M}$

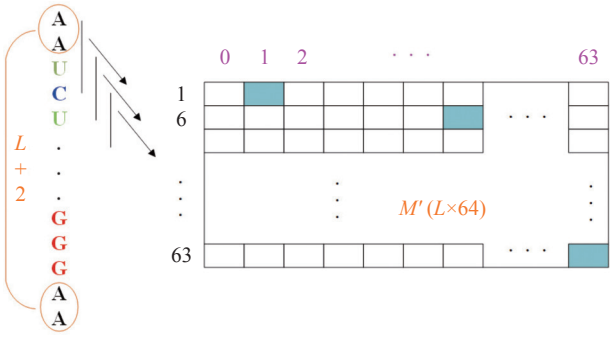

Figure 2 The encoding rules of loop-3-mer.

and $\boldsymbol{M'}$. In the data preprocessing stage, we perform the parallel processing on the two matrices separately to capture high-order latent features. Take $\boldsymbol{M}$ as an example, we first use CNN to capture the local features whose operation is as follows:

$$z_i = \boldsymbol{W}_i \otimes \boldsymbol{M}_i + b_i \tag{2}$$

where the weight matrix is $\boldsymbol{W}_i$, input is $\boldsymbol{M}_i$, and bias is b_i. $\boldsymbol{Z}$=$[z_1, z_2, z_3, \ldots, z_l]$ is the extracted feature map and there are l filters. After convolution, we introduce batch normalization (BN)-ReLU [30] to normalize $\boldsymbol{Z}$. ReLU may lead to the situation of the average value offset, resulting in difficulty convergence. BN-ReLU helps ReLU make better feature selection and speeds up the convergence. Through preprocessing, we can integrate two encoding matrices in feature-level rather than in data-level.

We get $\boldsymbol{R}$ and $\boldsymbol{R'}$ after data preprocessing on $\boldsymbol{M}$ and $\boldsymbol{M'}$, respectively. $\boldsymbol{R}$ and $\boldsymbol{R'}$ are then concatenated and fed into the encoder block.

4. Encoder-decoder architecture for feature learning

In circ-CNNED, we construct a CNN-based encoder-decoder architecture. We use TCN to construct the encoder block and CNN to construct the decoder block. The two blocks form a symmetrical structure (see Figure 1).

During data preprocessing, we obtain local features of circRNA sequences by CNN. In order to obtain the context-dependent information, most of the previous methods adopt the recurrent neural network (RNN). However, RNN will lose its ability as the distance increases due to vanishing gradient and runs for a long time. Thus, we introduce TCN [27] which considers the full temporal dependencies in circRNA sequences. Take filter_length=3 and *dilation*=[1, 2, 4, 8] as an example, the dilated acausal convolution is shown in Figure 3. Compared with CNN and RNN, TCN can perform dilated convolutions in parallel and residual connections. In this way, TCN has not only flexible receptive field size, but also stable gradients.

TCN is better at capturing temporal dependencies and trains faster than RNN. The receptive field can cover all values from the input sequence. The current, previous and future steps are cover as follows [26]:

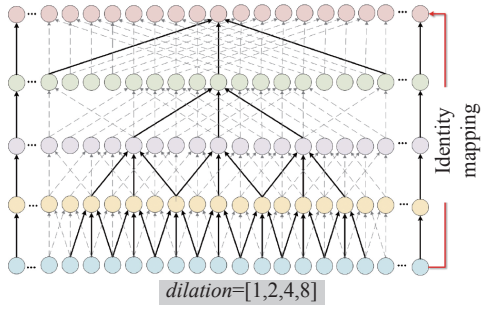

Figure 3 A dilated acausal convolution (filter_length=3).

$$\hat{\boldsymbol{C}}_t^l = f(\boldsymbol{W}^1 \boldsymbol{C}_{t-d}^{l-1} + \boldsymbol{W}^2 \boldsymbol{C}_t^{l-1} + \boldsymbol{W}^3 \boldsymbol{C}_{t+d}^{l-1} + \boldsymbol{b}) \tag{3}$$

where $\hat{\boldsymbol{C}}_t^l$ is the result at time t by acausal convolution. $\boldsymbol{C}^l$ gives the activations in the lth layer. Each layer includes a group of dilated convolutions with *dilation* and ReLU function f. The parameters are $\boldsymbol{W} = \{\boldsymbol{W}^1, \boldsymbol{W}^2, \boldsymbol{W}^3\}$ and the bias vector $\boldsymbol{b}$. Residual connection is used to eliminate vanishing or exploding gradients and to facilitate gradients flow. $\boldsymbol{C}^l$ is calculated as

$$\boldsymbol{C}_t^l = \boldsymbol{C}_t^{l-1} + \boldsymbol{V}\hat{\boldsymbol{C}}_t^l + e \tag{4}$$

where $\boldsymbol{V}$ is a group of weights and e is bias in residual connection. Receptive field is proportional to dilated convolution, which conduces to longer memory. The parameter numbers are not increased while avoiding overfitting [27]. The receptive field at lth layer is

$$R_f = 1 + (L_f - 1)(2^l - 1) \tag{5}$$

where R_f is the receptive field. L_f is the length of filter.

In encoder block, we adopt four TCN layers. Each layer has 64 filters and the filter_size is 10. After TCN, a max-pooling layer is used as down-sampling and we get the output of encoder block through it. We then feed the output of the encoder block to the up-sampling layer. We use four CNN layers whose computing mechanism is shown as equation (2) to construct a decoder block. The output of the last CNN layer is the output of the decoder block which is used for binding site prediction.

5. Global expectation pooling for classification

In classification block, we introduce a global expectation pooling layer to assign weights to features and a FC-softmax structure to classify.

Global expectation pooling [31] contains two sublayers (see Figure 4): 1) An one-dimensional (1D) global max-pooling. 2) A dense layer without parameterized weights of size 1. It calculates the weighted averages of max pooled values. In this way, there are no additional parameters.

Suppose that jth feature is the input of global expectation pooling: $Input_j$=$\boldsymbol{Q}_i^j$, where $i = 1, 2, \ldots, L$. The mathematical formulas are shown below.

$$\text{First sub-layer: } \boldsymbol{B}_j = \max_{i=1,2,\ldots L}\{\boldsymbol{Q}_i^j\} \tag{6}$$

$$\text{Second sub-layer: } output = \sum_{i}^{L} prob \times (\boldsymbol{Q}_i^j)^{\mathrm{T}} \quad (7)$$

$$prob = \exp\left\{m(\boldsymbol{Q}_i^j - \boldsymbol{B}_j)\right\} / \sum \exp\{m(\boldsymbol{Q}_i^j - \boldsymbol{B}_j)\} \quad (8)$$

where $m > 0$ and in circ-CNNED, we set m=1. *prob* is the weight assignment. Larger input values tend to have larger weights.

After global expectation pooling layer, we utilize a FC-softmax structure to classify. Through it, two results between (0,1) represent the probability of two classes.

III. Results

1. Performance of circ-CNNED

In data preprocessing, we set the number of filters as 128. In encoder block, we set the number of filters as 64 and the pool_size is 5. In decoder block, we set the number of filters as 128. While training, we use Adam as optimizer and categorical_crossentropy as loss function. We set batch_size as 64 and max_epoches as 50.

We obtain receiver operating characteristic (ROC) curve and precision/recall (PR) curve of circ-CNNED across 37 datasets (see Figure 5). Besides, we also evaluate circ-CNNED using other four metrics: ACC, Precision, Recall and F1_Score. From Figure 5, the largest AUC is 0.9870 (AUF1) and the minimum is 0.7800 (ZC3H7B). The largest AUPR is 0.9853 (AUF1) and the minimum is 0.7649 (TNRC6). circCNNED is effective under a variety of evaluation metrics.

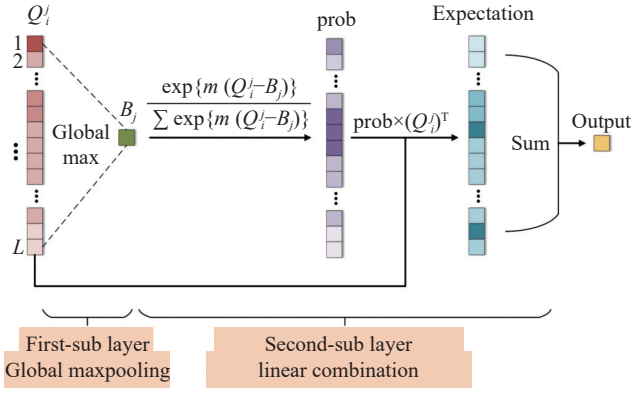

Figure 4 The architecture of global expectation pooling.

2. Comparison with other methods

We compare circ-CNNED with four existing prediction methods:

- CRIP [24]: Use a hybrid neural network and a stacked encoding scheme based on codon.
- PASSION [32]: Use the concatenated artificial neural network (ANN) and hybrid deep neural network.
- CRBPDL [33]: Adopt deep hierarchical network integrated with Adaboost.
- icircRBP-DHN [34]: Merge BiGRUs and a deep multi-scale residual network.

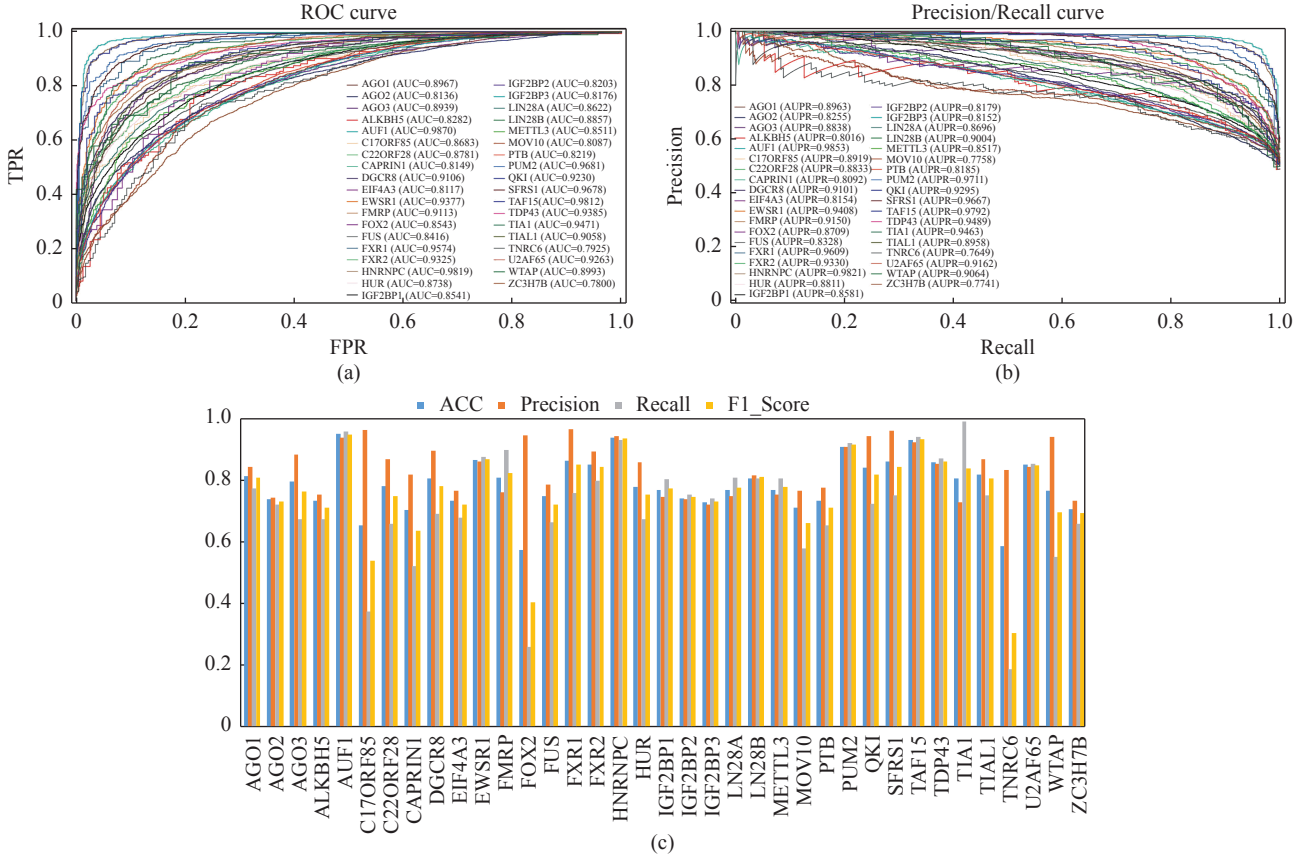

Figure 5 Results of circ-CNNED across 37 datasets. (a) The ROC curve; (b) The PR curve; (c) Other four metrics.

To ensure a fair comparison, all methods use the same datasets and their own hyperparameter settings. To compare these methods, we utilize AUC which has no fixed threshold. From Table 1 and Figure 6 (The line in the middle of each body represents the mean value), circ-CNNED is obviously superior than CRBPDL, CRIP and PASSION. The average, the maximum and the minimum AUC of circ-CNNED are all higher than them. The average AUC of circ-CNNED is near with icircRBP-DHN. However, the number of improved cases between them is 23/37 which indicates that circ-CNNED is better than icircRBP-DHN. Overall, circ-CNNED outperforms these four latest methods.

In order to verify whether the differences between circ-CNNED and these methods are significant, we do the t-test using AUC (see Figure 7).

Obviously, circ-CNNED is much better than CRIP and CRBPDL. Compared with PASSION and icircRBP-

Table 1 The comparisons of circ-CNNED with other methods using AUC

Datasets	circ-CNNED	PASSION	CRIP	CRBPDL	icircRBP-DHN
AGO1	0.8967	0.9050	**0.9052**	0.8160	0.8992
AGO2	0.8136	**0.8210**	0.8062	0.7017	0.7940
AGO3	0.8939	**0.9090**	0.8911	0.8198	0.8622
ALKBH5	0.8282	0.7520	0.7865	**0.9325**	0.7909
AUF1	**0.9870**	0.9780	0.9820	0.9360	0.9769
C17ORF85	0.8683	0.8590	0.8006	0.8328	**0.9576**
C22ORF28	0.8781	**0.8920**	0.8590	0.8061	0.8915
CAPRIN1	0.8149	0.8580	0.8313	0.7674	**0.8669**
DGCR8	0.9106	**0.9160**	0.9028	0.8201	0.9055
EIF4A3	0.8117	**0.8230**	0.8084	0.7165	0.8017
EWSR1	0.9377	**0.9380**	0.9239	0.8664	0.9340
FMRP	**0.9113**	0.9000	0.8868	0.7922	0.8922
FOX2	0.8543	0.8300	0.7735	0.7422	**0.9531**
FUS	0.8416	**0.8570**	0.8418	0.7668	0.8569
FXR1	**0.9574**	0.9550	0.9269	0.9321	0.9065
FXR2	0.9325	**0.9400**	0.9143	0.8580	0.9338
HNRNPC	**0.9819**	0.9760	0.9710	0.9083	0.9732
HUR	0.8738	**0.8770**	0.8768	0.7693	0.8713
IGF2BP1	**0.8541**	0.8410	0.8436	0.7579	0.8504
IGF2BP2	0.8203	0.8270	0.8294	0.7416	**0.8328**
IGF2BP3	0.8176	**0.8310**	0.8170	0.7103	0.8187
LIN28A	0.8622	**0.8740**	0.8649	0.7582	0.8560
LIN28B	0.8857	**0.8890**	0.8662	0.7917	0.8849
METTL3	0.8511	**0.8710**	0.8057	0.7583	0.8203
MOV10	0.8087	0.8340	0.8379	0.7593	**0.8404**
PTB	0.8219	**0.8290**	0.8186	0.7322	0.8216
PUM2	**0.9681**	0.9480	0.9619	0.9194	0.9512
QKI	0.9230	0.9270	0.9132	**0.9313**	0.9039
SFRS1	**0.9678**	0.9640	0.9641	0.8877	0.9598
TAF15	**0.9812**	0.9670	0.9804	0.7808	0.8899
TDP43	**0.9385**	0.9330	0.9274	0.8489	0.9267
TIA1	**0.9471**	0.9330	0.9278	0.8939	0.9433
TIAL1	0.9058	**0.9060**	0.9028	0.8393	0.8396
TNRC6	0.7925	0.7760	0.7756	0.8206	**0.9635**
U2AF65	0.9263	**0.9300**	0.9183	0.8475	0.9196
WTAP	0.8993	0.7940	0.7344	0.5602	**0.9220**
ZC3H7B	0.7800	0.8040	0.7918	0.7172	**0.8041**

Note: Bold values are the maximum on the same dataset and the underlined values are the second maximum.

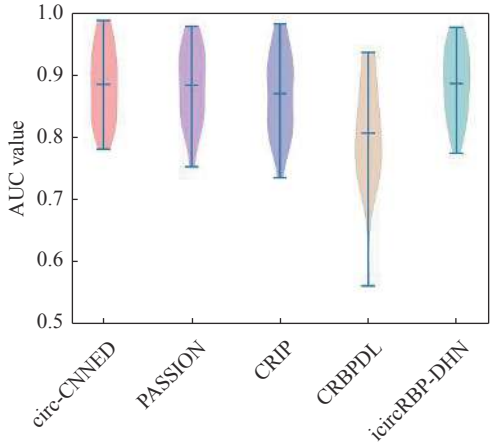

Figure 6 Performance comparison of different methods.

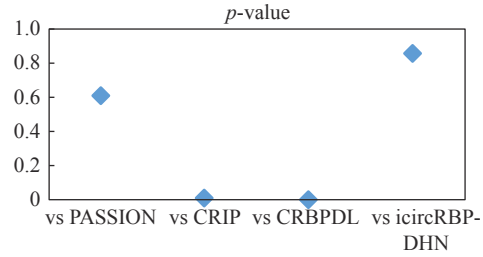

Figure 7 p-value between circ-CNNED with four comparison methods.

DHN, the performance of circCNNED is very close to them. In terms of the model structures of these comparison methods, circ-CNNED performs the preprocessing before concatenating the two encoding matrices which helps to get a better fusion. Besides, circ-CNNED uses TCN to learn context dependence which can avoid common defects of RNN.

3. Ablation experiments

1) Comparison of two encoding mechanisms

In circ-CNNED, we adopt two encoding mechanisms. To demonstrate the contribution of them, we build the following structures:

- circ-one: Only use the one-hot encoding mechanism.
- circ-loop: Only the loop-3-mer encoding mechanism is used.
- circ-CNNED: The structure proposed in this paper.

The comparison results can be seen in Figure 8 (The indigo line represents the median and the green triangle represents the mean) and Table 2. The performance of circ-CNNED is obviously better than the other two structures. The other two structures both contribute significantly to the performance of circ-CNNED. Compared to circ-loop, circ-one is superior which probably is because of the dimension of two matrices.

2) The effectiveness of global expectation pooling

To verify the effectiveness of global expectation pooling, we build the following structure:

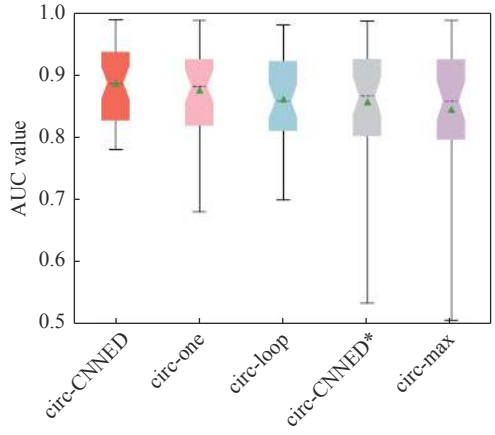

Figure 8 Performance comparison of different structures.

Table 2 The comparisons of different structures.

	Improved cases	Average improved
circ-CNNED vs. circ-one	27/37	1.2%
circ-CNNED vs. circ-loop	33/37	2.8%
circ-CNNED vs. circ-CNNED*	34/37	3.5%
circ-CNNED vs. circ-max	37/37	4.9%

- circ-max: Use max-pooling to replace global expectation pooling.

The comparison results can be seen in Figure 8, Table 2 and the performance of circ-CNNED is obviously better than circ-max. Compared with max-pooling, global expectation pooling is based on probability interpretability and helps us better understand the potential biological sequence from the perspective of statistical pattern.

3) The effectiveness of data preprocessing

In circ-CNNED, we preprocess the data before fusion. In this section, we remove the data preprocessing block (block "(A)" in Figure 1) to construct a new structure circ-CNNED*. The comparison results can be seen in Figures 8 and 9, Table 2. The experimental results show that the data preprocessing is of great importance. Through this, circ-CNNED gets more useful information.

4. Motif enrichment analysis

By analyzing the experimental results, we find that

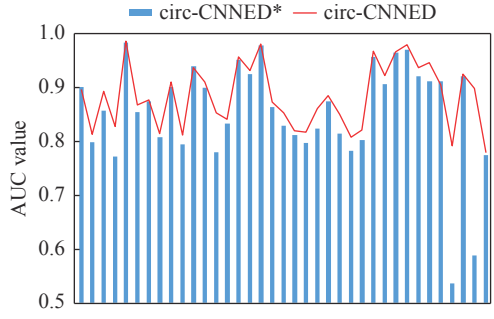

Figure 9 Performance comparison between circ-CNNED and circ-CNNED*.

the improvement of circ-CNNED is more pronounced on small datasets, especially in FXR1, PUM2, TAF15 and WTAP.

We utilize AME [35] and CentriMo [36] respectively to do motif enrichment analysis [37]. The results can be seen in Figures 10 and 11. Compared to AME, CentriMo is characterized by finding known motifs at fixed sites, i.e. motifs are in the same position in all sequences. The motifs enriched by CentriMo are concentrated on the center of the sequences and the results of AME and CentriMo are similar. As mentioned above, the center of sequences are binding sites. circ-CNNED is probably more sensitive to binding sites so that its performance on these four datasets improves significantly.

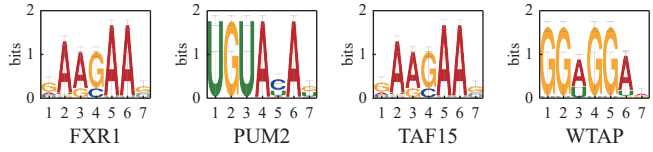

Figure 10 The motif enrichment results of AME.

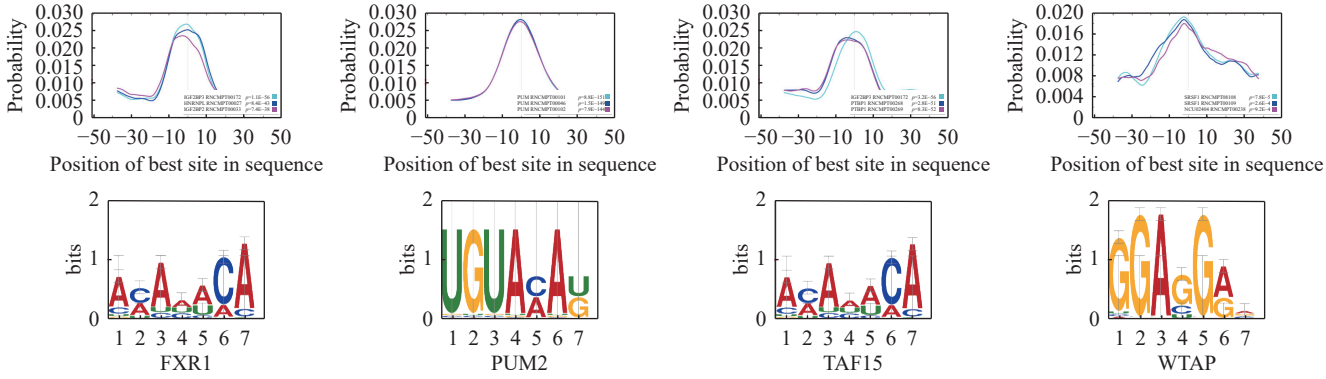

Figure 11 The motif enrichment results of CentriMo. The motifs shown have the smallest p-value.

IV. Conclusion

In this study, we come up with an original method, circ-CNNED, for prediction of RBP binding sites on circRNAs. We use one-hot and loop-3-mer respectively to get two raw matrices of circRNA sequences. Then we adopt CNN to preprocess them thus we can obtain useful information. Next, we utilize TCN to construct an encoder block and CNN to construct a decoder block. Through global expectation pooling, we weight the features and reduce their dimensionality. Finally we use a FC-softmax struture to classify. On 37 datasets, we compare circ-CNNED with four latest methods and perform ablation experiments. Experiment results demonstrate that circ-CNNED is effective for prediction of RBP binding sites on circRNAs.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant Nos. 62272288, 61972451, and U22A2041), the Fundamental Research Funds for the Central Universities, and Shenzhen Key Laboratory of Intelligent Bioinformatics (Grant No. ZDSYS20220422103800001).

References

[1] C. X. Liu and L. L. Chen, "Circular RNAs: Characterization, cellular roles, and applications," *Cell*, vol. 185, no. 13, article no. 2390, 2022.

[2] Z. Y. Wu, X. Yu, S. J. Zhang, *et al.*, "Mechanism underlying circRNA dysregulation in the TME of digestive system cancer," *Frontiers in Immunology*, vol. 13, article no. 951561, 2022.

[3] X. X. Zeng, Y. Zhong, W. Lin, *et al.*, "Predicting disease-associated circular RNAs using deep forests combined with positive-unlabeled learning methods," *Briefings in Bioinformatics*, vol. 21, no. 4, pp. 1425–1436, 2020.

[4] Y. Pan, X. J. Lei, and Y. C. Zhang, "Association predictions of genomics, proteinomics, transcriptomics, microbiome, metabolomics, pathomics, radiomics, drug, symptoms, environment factor, and disease networks: A comprehensive approach," *Medicinal Research Reviews*, vol. 42, no. 1, pp. 441–461, 2022.

[5] C. C. Wang, C. D. Han, Q. Zhao, *et al.*, "Circular RNAs and complex diseases: From experimental results to computational models," *Briefings in Bioinformatics*, vol. 22, no. 6, article no. bbab286, 2021.

[6] L. S. Kristensen, T. Jakobsen, H. Hager, *et al.*, "The emerging roles of circRNAs in cancer and oncology," *Nature reviews Clinical oncology*, vol. 19, no. 3, pp. 188–206, 2022.

[7] X. Dong, K. Chen, W. B. Chen, *et al.*, "circRIP: An accurate tool for identifying circRNA-RBP interactions," *Briefings in Bioinformatics*, vol. 23, no. 4, article no. bbac186, 2022.

[8] W. W. Du, C. Zhang, W. N. Yang, *et al.*, "Identifying and characterizing circRNA-protein interaction," *Theranostics*, vol. 7, no. 17, pp. 4183–4191, 2017.

[9] W. W. Du, W. N. Yang, E. Liu, *et al.*, "Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2," *Nucleic Acids Research*, vol. 44, no. 6, pp. 2846–2858, 2016.

[10] S. B. Qu, X. S. Yang, X. L. Li, *et al.*, "Circular RNA: A new star of noncoding RNAs," *Cancer Letters*, vol. 365, no. 2, pp. 141–148, 2015.

[11] M. Gagliardi and M. R. Matarazzo, "RIP: RNA immunoprecipitation," in *Polycomb Group Proteins*, C. Lanzuolo and B. Bodega, Eds. Humana Press, New York, NY, USA, pp. 73–86, 2016.

[12] C. Barnes and A. Kanhere, "Identification of RNA-protein interactions through in vitro RNA pull-down assays," in *Polycomb Group Proteins*, C. Lanzuolo and B. Bodega, Eds. Humana Press, New York, NY, USA, pp. 99–113, 2016.

[13] D. D. Licatalosi, A. Mele, J. J. Fak, *et al.*, "HITS-CLIP yields genome-wide insights into brain alternative RNA pro-

cessing," *Nature*, vol. 456, no. 7221, pp. 464–469, 2008.
[14] P. Glažar, P. Papavasileiou, and N. Rajewsky, "circBase: A database for circular RNAs," *RNA*, vol. 20, no. 11, pp. 1666–1670, 2014.
[15] D. B. Dudekula, A. C. Panda, I. Grammatikakis, *et al.*, "CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs," *RNA Biology*, vol. 13, no. 1, pp. 34–42, 2016.
[16] S. Ghosal, S. Das, R. Sen, *et al.*, "Circ2Traits: A comprehensive database for circular RNA potentially associated with disease and traits," *Frontiers in Genetics*, vol. 4, article no. 283, 2013.
[17] X. P. Chen, P. Han, T. Zhou, *et al.*, "circRNADb: A comprehensive database for human circular RNAs with protein-coding annotations," *Scientific Reports*, vol. 6, no. 1, article no. 34985, 2016.
[18] M. Liu, Q. Wang, J. Shen, *et al.*, "Circbank: A comprehensive database for circRNA with standard nomenclature," *RNA Biology*, vol. 16, no. 7, pp. 899–905, 2019.
[19] Y. J. Guo, X. J. Lei, L. Liu, *et al.*, "circ2CBA: Prediction of circRNA-RBP binding sites combining deep learning and attention mechanism," *Frontiers of Computer Science*, vol. 17, article no. 175904, 2022.
[20] P. P. Li and Z. P. Liu, "PST-PRNA: Prediction of RNA-binding sites using protein surface topography and deep learning," *Bioinformatics*, vol. 38, no. 8, pp. 2162–2168, 2022.
[21] X. J. Lei, T. B. Mudiyanselage, Y. C. Zhang, *et al.*, "A comprehensive survey on computational methods of non-coding RNA and disease association prediction," *Briefings in bioinformatics* , vol. 22, no. 4, article no. bbaa350, 2021.
[22] M. T. Niu, Q. Zou, and C. Y. Wang, "GMNN2CD: Identification of circRNA-disease associations based on variational inference and graph Markov neural networks," *Bioinformatics*, vol. 38, no. 8, pp. 2246–2253, 2022.
[23] Z. F. Wang and X. J. Lei, "Identifying the sequence specificities of circRNA-binding proteins based on a capsule network architecture," *BMC Bioinformatics*, vol. 22, no. 1, article no. 19, 2021.
[24] K. M. Zhang, X. Y. Pan, Y. Yang, *et al.*, "CRIP: Predicting circRNA-RBP-binding sites using a codon-based encoding and hybrid deep neural networks," *RNA*, vol. 25, no. 12, pp. 1604–1615, 2019.
[25] Z. F. Wang and X. J. Lei, "Prediction of RBP binding sites on circRNAs using an LSTM-based deep sequence learning architecture," *Briefings in Bioinformatics*, vol. 22, no. 6, article no. bbab342, 2021.
[26] Y. N. Yang, Z. L. Hou, Y. S. Wang, *et al.*, "HCRNet: High-throughput circRNA-binding event identification from CLIP-seq data using deep temporal convolutional network," *Briefings in Bioinformatics*, vol. 23, no. 2, article no. bbac027, 2022.
[27] S. J. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint*, arXiv: 1803.01271, 2018.
[28] Z. F. Wang and X. J. Lei, "A web server for identifying circRNA-RBP variable-length binding sites based on stacked generalization ensemble deep learning network," *Methods*, vol. 205, pp. 179–190, 2022.
[29] Y. Huang, B. F. Niu, Y. Gao, *et al.*, "CD-HIT Suite: A web server for clustering and comparing biological sequences," *Bioinformatics*, vol. 26, no. 5, pp. 680–682, 2010.
[30] Y. J. Guo and X. J. Lei, "A pseudo-Siamese framework for circRNA-RBP binding sites prediction integrating BiLSTM and soft attention mechanism," *Methods*, vol. 207, pp. 57–64, 2022.
[31] X. Luo, X. M. Tu, Y. Ding, *et al.*, "Expectation pooling: An effective and interpretable pooling method for predicting DNA-protein binding," *Bioinformatics*, vol. 36, no. 5, pp. 1405–1412, 2020.
[32] C. Z. Jia, Y. Bi, J. X. Chen, *et al.*, "PASSION: An ensemble neural network approach for identifying the binding sites of RBPs on circRNAs," *Bioinformatics*, vol. 36, no. 15, pp. 4276–4282, 2020.
[33] M. T. Niu, Q. Zou, and C. Lin, "CRBPDL: Identification of circRNA-RBP interaction sites using an ensemble neural network approach," *PLoS Computational Biology*, vol. 18, no. 1, article no. e1009798, 2022.
[34] Y. N. Yang, Z. L. Hou, Z. Q. Ma, *et al.*, "iCircRBP-DHN: Identification of circRNA-RBP interaction sites using deep hierarchical network," *Briefings in Bioinformatics*, vol. 22, no. 4, article no. bbaa274, 2021.
[35] R. C. McLeay and T. L. Bailey, "Motif Enrichment Analysis: A unified framework and an evaluation on ChIP data," *BMC Bioinformatics*, vol. 11, article no. 165, 2010.
[36] T. L. Bailey and P. Machanick, "Inferring direct DNA binding from ChIP-seq," *Nucleic Acids Research*, vol. 40, no. 17, article no. e128, 2012.
[37] P. Machanick and T. L. Bailey, "MEME-ChIP: Motif analysis of large DNA datasets," *Bioinformatics*, vol. 27, no. 12, pp. 1696–1697, 2011.

Yajing GUO is currently pursuing the Ph.D. degree in the School of Computer Science at Shaanxi Normal University, Xi'an, China. Her research interests include bioinformatics and deep learning.
(Email: guoyajing@snnu.edu.cn)

Xiujuan LEI is a Professor in the School of Computer Science at Shaanxi Normal University, Xi'an, China. Her research interests include intelligent computing and bioinformatics. (Email: xjlei@snnu.edu.cn)

Yi PAN is currently a Professor of the Faculty of Computer Science and Control Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China. He has served as Chair of Computer Science Department at Georgia State University during 2005–2020. His current research interests mainly include bioinformatics and health informatics using big data analytics, cloud computing, and machine learning technologies.
(Email:yi.pan@siat.ac.cn)