Communications in Information and Systems
Volume 22, Number 3, 363–382, 2022

A combined approach to RNA secondary structure prediction based on deep learning and minimum free energy model

Xiaoling He*, Xiujuan Ou*, Hong Yao*, Jun Wang, and Yi Xiao†

The secondary structures of RNAs are the basis of building their tertiary structures and understanding their functions. Many methods of RNA secondary structure prediction have been developed and can be divided into single-sequence and multiple-sequence methods, depending on one sequence or multiple sequences as input. Here we present a method, called 2dRNAx, that combines multiple-sequences method with single-sequence method. The results show that this combined method gives significantly higher accuracy than current multiple-sequences methods. Current version of the method only predicts canonical base pairs without pseudoknots, lone pairs and multiplets.

1. Introduction

The secondary structures of RNAs are defined as special patterns of A-U, G-C and G-U base pairings. They are important to their biological functions [1] and tertiary structure [2, 3, 4, 5, 6]. Therefore, many methods have been developed to predict the secondary structures. These methods can be divided into single-sequence based [7, 8] or multiple-sequences based ones [9].

The single-sequence methods usually adopt the minimum free energy (MFE) principle [7, 8, 10, 11, 12, 13, 14]. The advantage of these methods is that only the sequence of target RNA is needed as their inputs. However, since accurate calculation of free energy is difficult, especially for the loop regions of RNAs, the prediction accuracies of these methods are only about 70% [15, 16]. The multiple-sequence methods use the evolution information of the homologous sequences of the target RNA to infer their common secondary structures and their prediction accuracies are about 70–80%, usually higher than those of the single-sequence methods [6, 9, 17, 18, 19, 20, 21, 22, 23, 24]. Since these methods need aligned

*Who made equal contributions.
†Cooresponding author.

homologous sequences of the target RNA as input, they will give lower prediction accuracy or not work if the number of the homologous sequences are not enough. This suggests that we can combine the advantages of the single-sequence and multiple-sequences methods to improve accuracy of RNA secondary structure prediction.

It has been shown that the prediction accuracy of the single-sequence MFE methods could be improved by using additional constraints of base pairing provided by experiment, like SHAPE data [25, 26]. On the other hand, it was shown that the prediction accuracy of the multiple-sequence methods could be improved by using machine learning methods [27]. Motivated by these, here we propose a two-stages method to combine the single-sequence and multiple-sequences methods: For a target RNA, firstly we use a deep learning (DL) model to predict the possible base pairs (A-U, C-G and G-U) based on its homologous sequences and secondly, we use the MFE methods to do further prediction by using the predicted base pairs by the DL model as constraints. It is noted that the first step itself forms a multiple-sequences method of RNA secondary structure prediction based on deep learning approach. Comparing with the previous methods, this combined method gives higher accuracy than current multiple-sequence methods.

2. Methods and materials

2.1. Prediction workflow of 2dRNAx

Figure 1 shows the prediction workflow of our combined method, called 2dRNAx. For a target RNA sequence, we firstly find its multiple sequence alignment (MSA) from Rfam [28] database or using other methods. The pairwise frequencies of the bases in the MSA (see the following for the details) are the input of the trained deep learning model to predict the pairing probability of any two bases. Two bases are considered to form base pair if their pairing probability is larger than a threshold (0.5 in this work). These base pairs are further treated by a remove-and-expand procedure (see the following for the details) and then taken as constraints to input to an MFE algorithm to do further prediction. The results of the MFE method are also treated by the same removing procedure above to remove lone pairs and then taken as the final prediction output of the combined method.

2.2. Deep learning (DL) neural network

Our DL model is a U-net convolutional neural network model. We choose the U-net based on two reasons: first, the secondary structure of an RNA

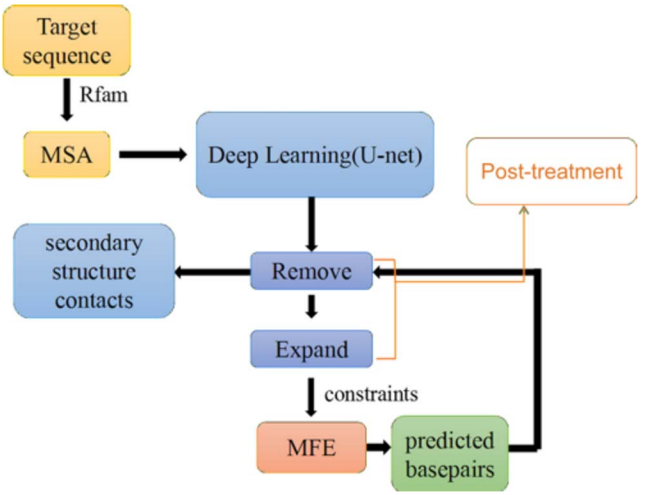

Figure 1: The prediction workflow of the combined method 2dRNAx.

can be represented by a two-dimensional matrix, i.e., a dot image; second, the samples of RNA secondary structures are not very large. The U-net has good performance in image segmentation and it is also good for small samples [29].

The U-net used in this work contains an input layer, nine hidden layers, and an output layer as usual [29] but with two modifications: (1) each convolution layer is followed by a Batch normalization layer to accelerate the model convergence and (2) a focal loss is used to solve the inharmonious problem of positive (pairing bases) and negative (non-pairing bases) samples [30]. In the training process, we use a 4-fold cross validation to reliably evaluate the model. In the testing process, the intersection of the results of the four trained models is remained. The detailed structure of the model is shown in Figure 2 and the details of each layer are as follows:

1) The input layer: The size of the input feature map is $L \times L \times 25$, where the L of the first and second dimensions is the length of the target RNA sequence, and the third dimension is the arrays of pairwise base frequencies in aligned homologous sequences of the target RNA. For example, the element (i, j) of AU array is the frequency that the nucleotide A in the ith column and the nucleotide U in the jth column of the aligned homologous sequences occur simultaneously in one row or one sequence. Since there are five symbols (A, C, G, U and gap "-") in the aligned homologous sequences, we have 25 pairwise combinations

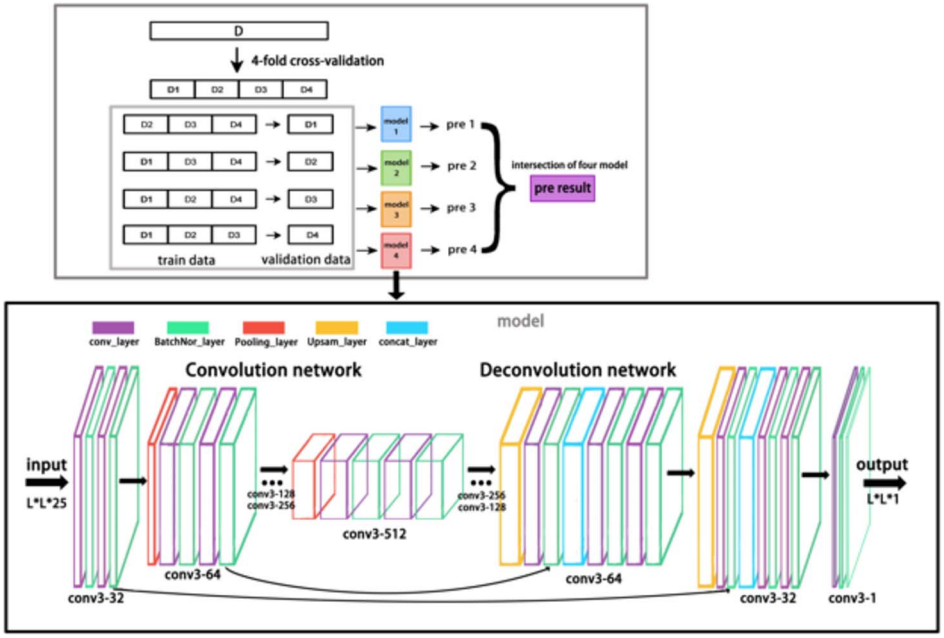

Figure 2: The layout of our improved U-net convolution neural network model [29]. L is the RNA sequence length. In training process, we use 4-fold cross validation to reliably evaluate model.

for them and so 25 arrays of the pairwise base frequency. Although we only focus on a very limited number of base pairs (AU, UA, CG, GC, GU, UG) in this work, in practice there are also base pairs other than them that are few but are very important to tertiary interactions. We hope our DL model can be expanded to predict these base pairs in future and so we have considered all 25 possibilities. Furthermore, considering all possible types of base pairs is easier for probability normalization. It is also noted that in current version of our DL model the lengths (L) of the target RNAs are limited to be not longer than 512nt and so the sequences that are shorter than 512nt should be padded with zeros.

2) The first layer: It includes two convolution layers, two Batch normalization layers, and one pooling layer. The number of convolution kernels of each convolution layer is 32. After the two convolution layers, the feature map size is changed from $L \times L \times 1$ to $L \times L \times 32$. The pooling layer size is 2×2, and the feature map size is changed from $L \times L \times 32$ to $L/2 \times L/2 \times 32$ through this pooling layer.

3) The second and third layers are similar in structure to the first layer except that the number of convolution kernels of the convolution layer is increased.
4) The fourth layer: In order to prevent over-fitting, a Dropout layer (the value is 0.5) is added behind the two pooling layers, and after the fourth layer, the feature map size becomes $L/16 \times L/16 \times 256$.
5) The fifth layer: this layer also includes a Dropout layer (the value is 0.5). Since there is no pooling layer, the size of the data will not shrink and the shape of the fifth layer of the data becomes $L/16 \times L/16 \times 512$.
6) The sixth layer: it includes the UpSampling2D layer, one convolution layer, one concatenate layer and two convolution layers; UpSampling factor is 2×2, the data is doubled, and the feature map size becomes $L/8 \times L/8 \times 512$; The concatenate layer connects the feature map of the Conv6_1 layer with the feature map of the Dropout4 layer according to the specified axis (here, the third dimension).
7) The seventh, eighth, and ninth layers are similar to the sixth layer. After the ninth layer, the feature map size becomes $L \times L \times 32$.
8) The output layer: It contains two convolution layers. For the first convolution layer, the number of convolution kernels is 2, the convolution kernel size is 3×3; ReLU is the activation function and the feature map size becomes $L \times L \times 2$; For the second convolution layer, the number of convolution kernels is 1 and the convolution kernel size is 1×1, Sigmod is the activation function, and the output feature map size is $L \times L \times 1$, where the first and second dimensions are the length of the RNA sequence, and the third dimension reflects the probability of each pair of bases that can form a base pair, being between 0 and 1.
9) For all convolutional layers from the first to the ninth layers, each convolution layer is followed by a Batch normalization layer, ReLU is the activation function; The convolution kernel size is 3×3, the step size is 1; The padding (filling mode) is "Same" (the input and output shapes are the same); The initializer of each layer weight matrix is "he_normal".

It is noted that in convolutional neural network, 32 is a HyperParameter (the number of filters computed by the convolution, every dimension is a feature). This value is needed to try. If the value is too small, some features can't be learned, resulting in underfitting, otherwise overfitting. In the U-net, the depth of the feature maps progressively increases (from 32 to 512) in convolution network and the depth of the feature maps progressively decreases (from 512 to 32) in deconvolution network [29].

2.3. Minimum free energy (MFE) method

For a target RNA, if only its RNA sequence is given, the minimum free energy principle is generally used to predict its secondary structure [15, 31, 32]. Our DL method uses the homologous sequences of a target sequence to predict all possible base pairs. In principle, it can give the information of all base pairs in the most stable structure, or more accurately in the native structure, but in practice it can only predict part of the base pairs. This is because that the quality of the homologous sequences is not the same for each target sequence, e.g., for some target sequences the number of the homologous sequences is very few or the aligned homologous sequences have a lot of gaps in some regions. This will lead to that the DL method cannot predict all standard base pairs in the native structure of these target sequences in some cases. To solve this problem, we use an MFE method to predict the remaining unpaired regions with the base pairs predicted by the DL model as constraints.

Many excellent algorithms of RNA secondary structure prediction based on the MFE principle are available. Since we need to input the predicted base pairs from the DL model as constraints into the algorithms, we will use the algorithms Mfold [8], RNAfold [11, 14] and RNAstructure [12].

2.4. Remove-and-expand post-treatment

In this work all the predictions by our DL and combined methods will be post-treated by a remove-and-expand procedure (Figure 1). This procedure is proposed simply based on the following reasons: (1) The RNA secondary structures in most current prediction methods are defined by the standard base pairs C-G, A-U and G-U since their binding interactions are much stronger than those of nonstandard base pairs. So, we only consider the standard base pairs too and remove the nonstandard base pairs. (2) From point of view of chemistry, one base can form stable base pair only with another one but not many bases and so we remove one-to-many base pairs predicted by DL. Furthermore, a lone base pair usually is not stable and so we also remove it. But this does not mean that one-to-many or lone base pairs do not exist in practice, and we neglect them since they are higher order interactions. (3) The longer the stems and the more stable and so we expand the stems as longer as possible. Of course, the expanded base pairs are limited only to the standard base pairs.

According to the reasons above, we only predict the secondary structures formed by the standard base pairs A-U, C-G and G-U and without

pseudoknots, we remove non-standard base pairs, pseudoknot base pairs, lone base pairs and one-to-many base pairs predicted from the DL model and the MFE methods. We'll also maximize the number of the base pairs A-U, C-G and G-U for the DL predictions. We do this in following steps:

1) Removing all base pairs other than A-U, C-G, and G-U in the predicted base pairs.
2) Removing all the predicted lone-pairs that cannot be extended to form consecutive canonical base pairs A-U, C-G and the wobble base pair G-U.
3) Removing the pairs in the predicted one-to-many pairs that have no or cannot be extended to form consecutive canonical base pairs A-U, C-G and the wobble base pair G-U.
4) Expanding or maximizing the base pairs A-U, C-G and G-U from the unpaired bases if they can form a continuous stacking with the predicted base pairs. During the expanding process, the hairpin loops are kept having at least four bases.
5) Removing the pseudoknot base pairs after expanding process.

2.5. Prediction performance estimation

Precision (PPV), sensitivity (STY) and Matthews correlation coefficients (MCC) [33, 34] are used to measure the performance of our method. They are defined as follows:

$$PPV = \frac{TP}{TP + FP},\ STY = \frac{TP}{TP + FN}$$

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (FP + TN) * (TN + FN) * (FN + TP)}}$$

where TP denotes true positive; FP, false positive; TN, true negative; FN, false negative.

These three performance evaluation metrics were usually used in the evaluation of RNA secondary structure prediction methods, and we use them because we want to compare with other methods. Furthermore, it is also for considering different aspects of the performance to use all these three metrics. You can just use one of the metrics, e.g., PPV if you focus on the precision. If you want to obtain a balance of PPV and STY, you can use MCC.

2.6. Training and testing sets

The RNAstrand dataset and PDB dataset in compaRNA [9] were frequently used as a training set and a testing set for multiple-sequence methods and they contain 1987 and 121 sequences, including their secondary structures and aligned homologous sequences, respectively. They can be downloaded from the website: http://iimcb.genesilico.pl/comparna/ [9]. The secondary structures of the RNAs in the PDB dataset were calculated from their 3D structures by RNAView [35]. We only selected the sequences with length less than 500 and so the training set finally contains 1128 sequences with the sequence lengths from 40 to 499 nucleotides. In order to find out which sequences in the testing set are similar to the training set, we use the tools provided by blast. The specific method is to use the training set as the database for blast search, and the testing set as the search data. If the similarity between the sequences in the training set and the sequences in the testing set is greater than 85%, then remove these sequences in the testing set to ensure that the sequences in the test set are not similar to those in the training set. Finally, the testing set contains 92 RNAs with the sequence lengths from 29 to 377 nucleotides. The number of the homologous sequences of the RNAs in these two datasets is from 6 to 1023.

We use a K-fold cross validation to train the U-net neural network model. Since our training set is not large, we choose K= 4. In fact, we found that the results were similar for K = 3, 4 and 5. Furthermore, the available data is randomly split into 4 partitions and so the training result should not depend on the initial split of the data. During the training, each of the four parts is taken as a validation set and the other three are as the training set.

3. Results

In the following we will evaluate the performance of the DL model only and the combined methods over the testing set, respectively. For the predictions of the DL model, if the predicted pairing probability of two bases by all the four trained models is larger than or equal to the threshold value 0.5, they are considered to form a base pair. We have also examined other threshold values but 0.5 gives the best performance. The results are shown in Table 1 and Figure 3. It is noted that the results have already been treated by the remove-and-expand procedure.

To compare with other multiple-sequence methods, CentroidAlifold [21], MXSCARNA [23] are selected because they are the two of the best multiple-sequence methods [9]. Other two [9] are RNAalifold [24] and Turbofold [22]

Table 1: The performances of different prediction methods

Methods	Number of RNAs	PPV	STY	MCC
DL	92	0.86	0.80	0.81
DL+Mfold	92	0.82	0.90	0.86
DL+RNAfold	92	0.82	0.90	0.86
DL+RNAstructure	92	0.82	0.91	0.86
Mfold	92	0.68	0.75	0.71
RNAfold	92	0.68	0.77	0.72
RNAstructure	92	0.69	0.78	0.72
CentroidAlifold*	87	0.83	0.48	0.61
MXSCARNA*	87	0.76	0.72	0.73

* The PPV, STY and MCC of CentroidAlifold and MXSCARNA are calculated based on the prediction results given at the websites: http://iimcb.genesilico.pl/comparna/atlas/pdb/. The number of RNAs for a method is that the method can predict at least one base pair for each of these RNAs.

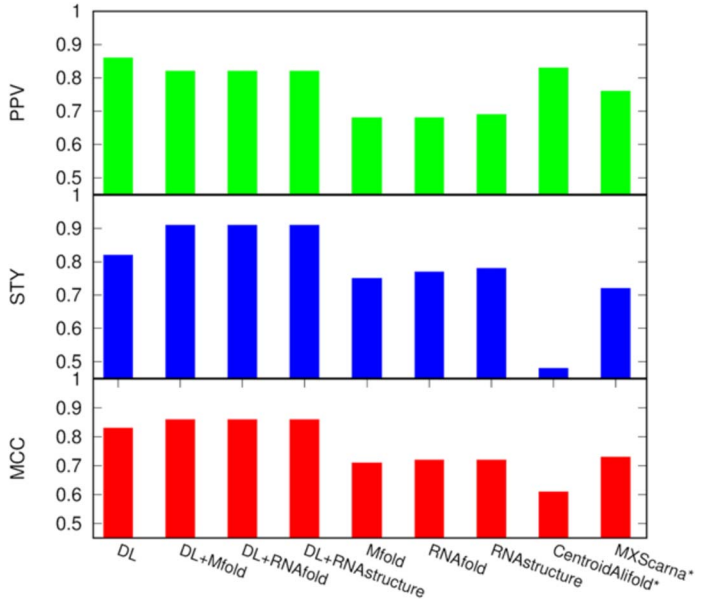

Figure 3: Performance (PPV colored in green, STY colored in blue, MCC colored in red) of different methods listed in Table 1.

but they can give prediction results only for 50 and 28 out of the 93 RNAs in the testing set, respectively, i.e., they do not predict any base pairs for other RNAs. So, it is not suitable to compare them with other methods. In com-

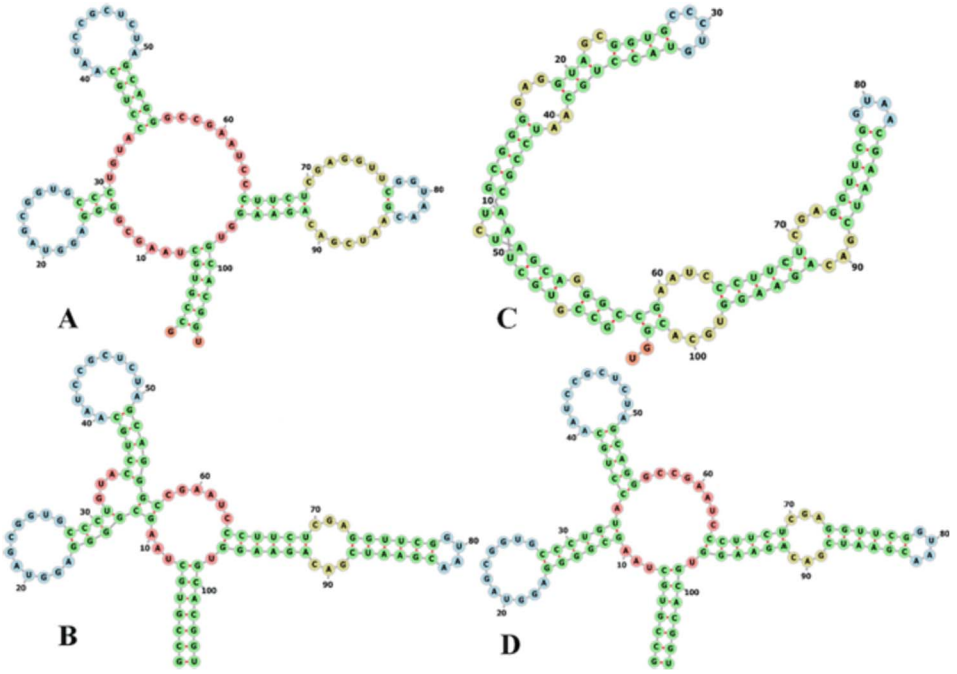

Figure 4: Prediction results for a RNA (PDB index: 4WFL_A). A: DL method; B: Native structure; C: RNAstructure; D: DL+ RNAstructure. All the secondary structures were drawn by the Forna [36].

parison with CentroidAlifold and MXSCARNA (see Table 1 and Figure 3), the MCCs and STYs of the combined methods are significantly higher. CentroidAlifold has slightly higher PPV (1%) than the combined methods but its STY and so MCC are much lower and are the lowest among all the methods. It is noted that the performance of the DL method only is significantly higher than those of CentroidAlifold and MXSCARNA on average.

Figures 4, 5, 6 show three examples of the predictions of different methods. We can see that the structure predicted by the DL model has long loops and that by RNAstructure is different from the native one in some cases while the predicted structure of the combined method is very closer to the native one.

4. Discussion

The results above show that in general the DL and combined methods have significantly better performance than the popular multiple-sequence meth-

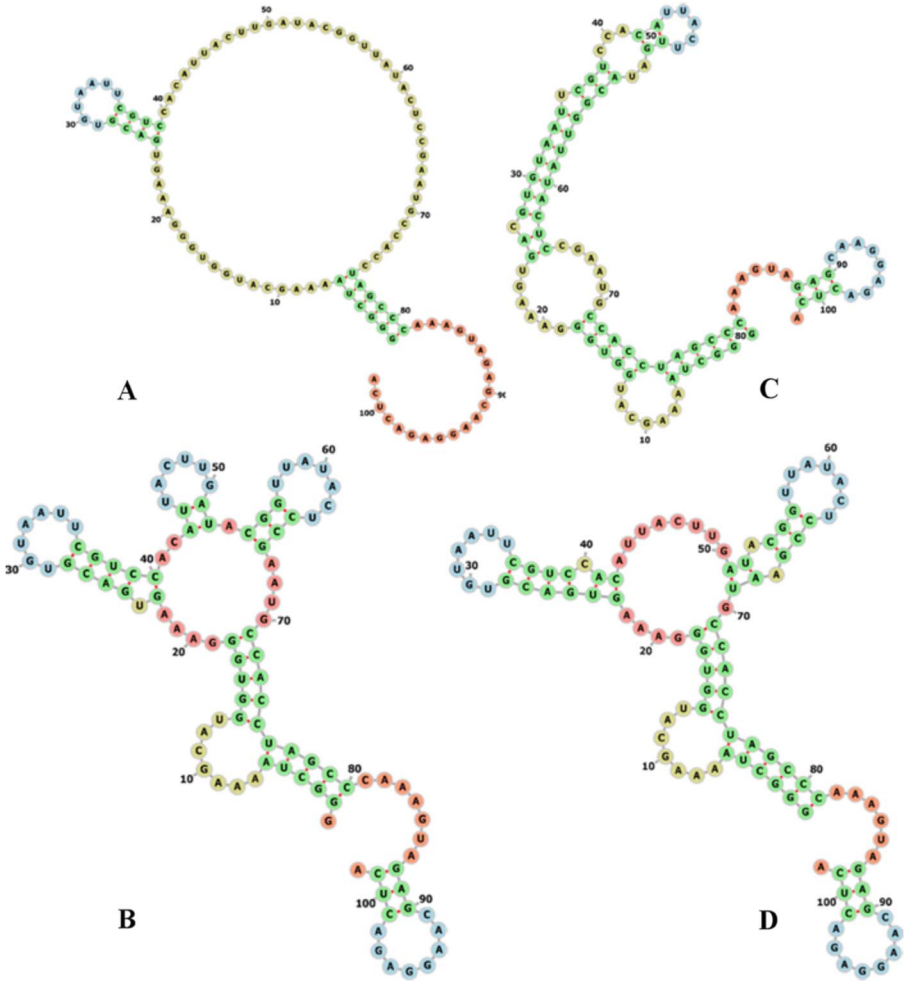

Figure 5: Prediction results for a RNA (PDB index: 4RFN_A). A: DL method; B: Native structure; C: RNAstructure; D: DL+ RNAstructure. All the secondary structures were drawn by the Forna [36].

ods. The DL model is good at increasing the precision (PPV) while the MFE model can further find more true positives or native base pairs (STY). On the other hand, the combined method can significantly increase the accuracy of the MFE methods if the homologous sequences of the target RNAs are available. It is noted that recently single-sequence methods of RNA secondary structure prediction using deep learning have been reported and they could also significantly increases the prediction precision (PPV) [37, 38, 39].

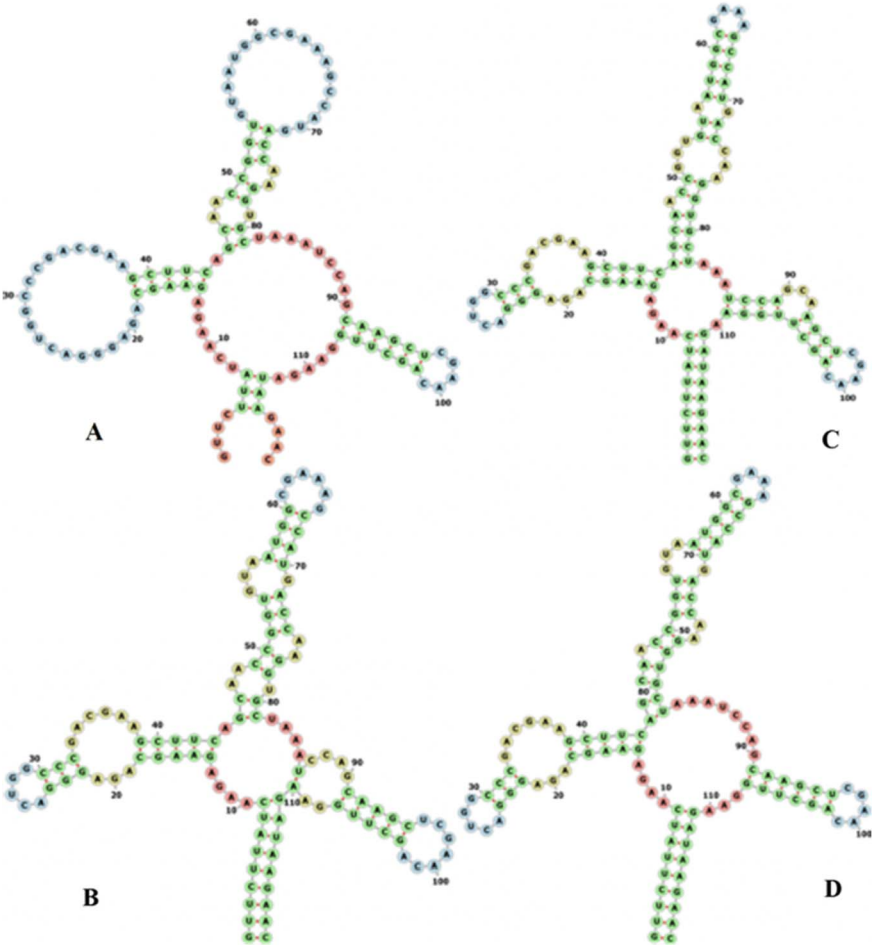

Figure 6: PPrediction results for an RNA (PDB index: 3NPB_A). A: DL method; B: Native structure; C: RNAstructure; D: DL+ RNAstructure. All the secondary structures were drawn by the Forna [36].

There was also a method of RNA secondary structure prediction that integrated the MFE method and the machine learning approach [40]. However, since these methods did not use information of homologous sequences as input, we'll not compare their results here.

Comparing with the DL method, the average PPVs of the combined methods decrease slightly but the average STYs increase significantly. To quantify this difference, we consider the ratio of the number of correctly predicted stems (successive base-pairing regions) to the total number of

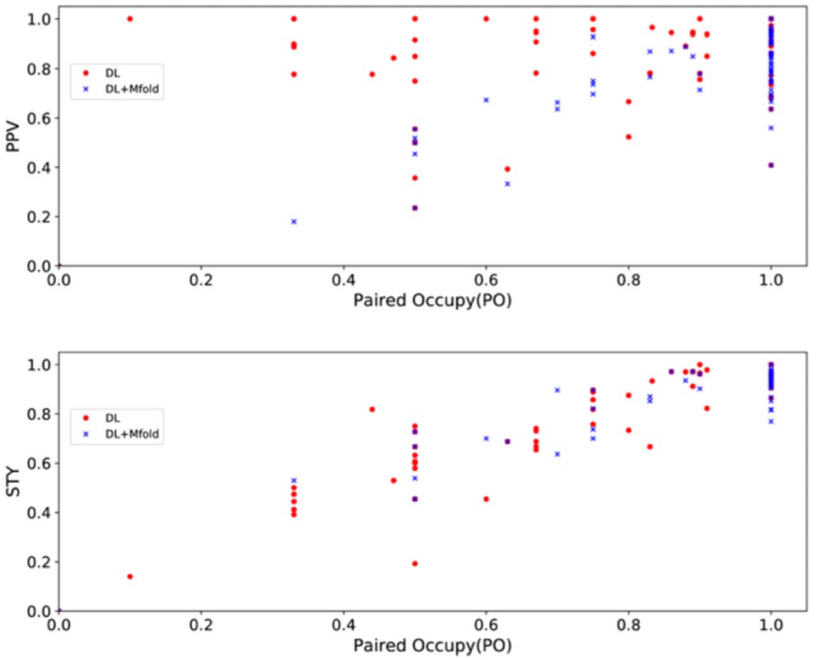

Figure 7: Paired occupy of the predictions of the DL and combined methods on the testing set.

stems in native structure for each RNA and we call it Paired Occupy (PO). For example, if the number of correctly predicted stems is 5, the total number of native stems is 10 in native structure, then the PO is $5/10 = 0.5$. A stem is considered as correctly predicted if at least one native base pair in it is predicted. For a stem, if too few native base pairs are predicted, usually other native base pairs can be found by the expanding step as shown above. The prediction results of the DL model without expanding are worse. If the predicted stem is longer than the actual one, it is considered as a correctly predicted one. It is noted that the definition here is used only for the Paired Occupy but not for PPV and STY.

The POs estimated by the DL and combined methods are shown in Figure 7, respectively. We find that many high PPV predictions by the DL model have low PO values, i.e., the predicted base pairs only distribute to a part of the stems in many cases although the DL method gives more predictions with high PPV. This may be the reason why the average PPV of the combined method is lower than that of the DL model since the MFE methods may introduce false positives. On the other hand, the predictions

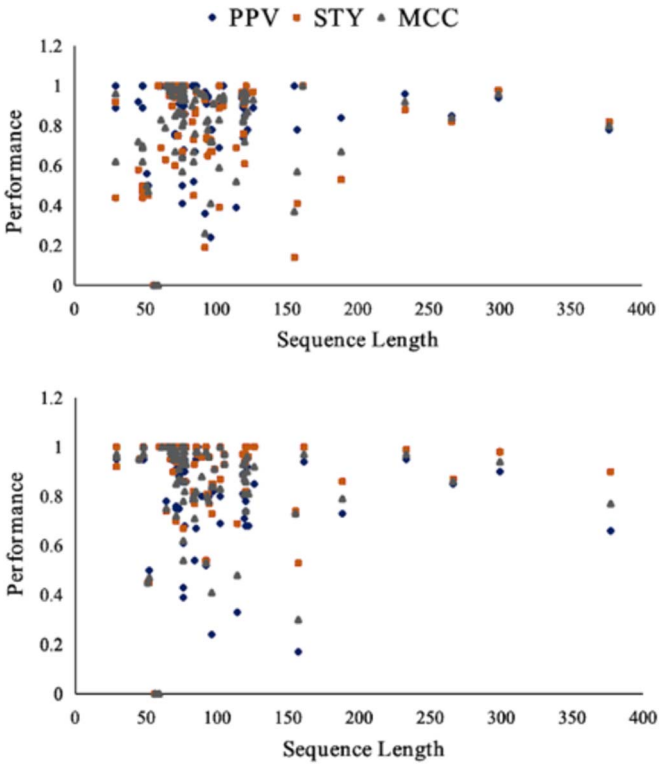

Figure 8: The scatter plot of performances of the DL (top) and combined (DL+RNAstrucutre) (bottom) methods with the sequence lengths (in nt) of the RNAs in the testing set.

by the combined method have a shift to the high PO values in comparison with those by the DL model, i.e., the predicted base pairs by the former cover more stems than those by the latter. This means that the predictions by the combined method have higher STY than those by the DL model, i.e., the MFE model can predict additional native base pairs on the basis of the DL prediction. This can also be seen from the lower plot in Figure 7 where the predictions by the combined method have a shift to both higher STY and PO values on average.

Figure 8 shows the performance of the DL and combined (DL+RNAstructure) methods with the sequence length of the RNAs in the testing set. We find that their performances have no clear dependence on the sequence lengths, at least when the lengths are shorter than 150nt. When the lengths are longer than 150nt, the performance of both methods are much better.

Table 2: Performances of our method for different types of RNAs in the testing set

RNA Type	Numbers	DL			combined(DL+RNAstructure)		
		PPV	STY	MCC	PPV	STY	MCC
Riboswitch	21	0.89	0.73	0.80	0.82	0.86	0.84
tRNA	34	0.87	0.96	0.91	0.86	0.97	0.91
rRNA	17	0.82	0.86	0.83	0.73	0.91	0.81
Ribozyme	2	0.46	0.29	0.36	0.47	0.47	0.47
tmRNA	1	0.78	0.82	0.80	0.66	0.90	0.77
Others	17	0.92	0.69	0.78	0.91	0.92	0.91

Table 3: RNA types and numbers in training and test sets

RNA Type	Numbers	
	Training Set	Test Set
Riboswitch	3	21
tRNA	208	34
Ribozyme	45	2
tmRNA	232	1
RNaseP	275	0
Intron gpI	15	0
rRNA	70	17
Telomerase	17	0
Vimentin	8	0
s2m	13	0
signal recognition particle	196	0
Others	46	18

However, this needs to be further validated since only a few RNAs have the lengths longer than 150nt in the testing set.

Table 2 shows the performances of the combined method (DL+RNAstructure) for different types of RNAs in the testing set (also see Table 3). The performances for tRNA and ribozyme are the highest and lowest, respectively. This may be due to the large or small number of them in the testing and training sets (Table 3). However, it is interesting to note that the performance of the DL and combined methods for riboswitch is higher than those for ribozyme and tmRNA although there are only three riboswitches in the training set. The reason for this needs further investigations.

It should be pointed out that the higher performance of 2dRNAx is only for the target RNAs that have enough homologous sequences. Otherwise, the DL model cannot work well and so 2dRNAx performs similarly as the MFE methods. As mentioned above, the performance of single-sequence method

can also be improved by using deep learning. So, it may be better to combine single-sequence and multiple-sequence methods by applying deep learning to both of them. It is also noted that the computational complexity of the combined method is similar to the MFE method since the DL model is time-consuming only in the training process.

In addition, it should be noted that the current method was trained and tested with Rfam prealigned sequences while Rfam may use experimentally determined secondary structures for their alignments. Thus, the performance may be different if other automatic alignment techniques are used.

5. Conclusion

In this paper, we combine deep learning with the traditional MFE methods to improve RNA secondary structure prediction. The results show that this combined method not only gives significantly higher accuracy than current multiple-sequence methods. The used DL model is good at increasing the precision (PPV) while the MFE methods are able to find more true positives or native base pairs (STY). Furthermore, merely the DL model already has better performance. These results show that the combination of single- and multiple-sequence methods may be a way of improving RNA secondary structure prediction.

6. Data and webserver availability

The data used for training and testing and 2dRNAx web server are available at http://biophy.hust.edu.cn/new/2dRNAx.

Acknowledgements

This work is supported by the NSFC under Grant No. 32071247.

References

[1] Y. Zhao, et al., *NONCODE 2016: an informative and valuable data source of long non-coding RNAs*. Nucleic Acids Res, 2016. **44**(D1): p. D203–D208.

[2] Y. J. Zhao, et al., *Automated and fast building of three-dimensional RNA structures*. Sci Rep, 2012. **2**: p. 734.

[3] J. Wang, et al., *3dRNAscore: a distance and torsion angle dependent evaluation function of 3D RNA structures*. Nucleic Acids Res, 2015. **43**(10): p. e63.

[4] J. Wang and Y. Xiao, *Using 3dRNA for RNA 3-D Structure Prediction and Evaluation*. Curr Proto Bioinformatics, 2017. **57**: p. 5.9.1–5.9.12.

[5] E. De Leonardis, et al., *RNA Secondary and Tertiary Structure Prediction by Tracing Nucleotide Co-Evolution with Direct Coupling Analysis*. Biophys J, 2016. **110**(3): p. 364a.

[6] E. De Leonardis, et al., *Direct-Coupling Analysis of nucleotide coevolution facilitates RNA secondary and tertiary structure prediction*. Nucleic Acids Res, 2015. **43**(21): p. 10444–10455.

[7] M. Zuker and P. Stiegler, *Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information*. Nucleic Acids Res, 1981. **9**(1): p. 133–148.

[8] M. Zuker, *Mfold web server for nucleic acid folding and hybridization prediction*. Nucleic Acids Res, 2003. **31**(13): p. 3406–3415.

[9] T. Puton, et al., *CompaRNA: a server for continuous benchmarking of automated methods for RNA secondary structure prediction*. Nucleic Acids Res, 2013. **41**(7): p. 4307–4323.

[10] M. Zuker, *On finding all suboptimal foldings of an RNA molecule*. Science, 1989. **244**(4900): p. 48–52.

[11] R. Lorenz, et al., *ViennaRNA Package 2.0*. Algorithms Mol Biol, 2011. **6**: p. 26.

[12] S. Bellaousov, et al., *RNAstructure: web servers for RNA secondary structure prediction and analysis*. Nucleic Acids Res, 2013. **41**(W1): p. W471–W474.

[13] S. Janssen and R. Giegerich, *The RNA shapes studio*. Bioinformatics, 2015. **31**(3): p. 423–425.

[14] A. R. Gruber, S.H. Bernhart, and R. Lorenz, *The ViennaRNA web services*. Methods Mol Biol, 2015. **1269**: p. 307–326.

[15] K. J. Doshi, et al., *Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction*. BMC Bioinformatics, 2004. **5**: p. 105.

[16] Y. Zhao, et al., *Evaluation of RNA secondary structure prediction for both base-pairing and topology*. Biophysics Reports, 2018. **4**(3): p. 123–132.

[17] Z. Tan, et al., *TurboFold II: RNA structural alignment and secondary structure prediction informed by multiple homologs*. Nucleic Acids Res, 2017. **45**(20): p. 11570–11581.

[18] F. L. Simonetti, et al., *MISTIC: Mutual information server to infer coevolution*. Nucleic Acids Res, 2013. **41**(W): p. W8–W14.

[19] X. L. He, et al., *Comparison of two algorithms of direct coupling analysis of protein*. Commun Inf Syst, 2019. **19**(1): p. 1–15.

[20] J. Wang, et al., *Optimization of RNA 3D structure prediction using evolutionary restraints of nucleotide-nucleotide interactions from direct coupling analysis*. Nucleic Acids Res, 2017. **45**(11): p. 6299–6309.

[21] M. Hamada, K. Sato, and K. Asai, *Improving the accuracy of predicting secondary structure for aligned RNA sequences*. Nucleic Acids Res, 2011. **39**(2): p. 393–402.

[22] A. O. Harmanci, G. Sharma, and D.H. Mathews, *TurboFold: iterative probabilistic estimation of secondary structures for multiple RNA sequences*. BMC Bioinformatics, 2011. **12**: p. 108.

[23] Y. Tabei, et al., *A fast structural multiple alignment method for long RNA sequences*. BMC Bioinformatics, 2008. **9**: p. 33.

[24] S. H. Bernhart, et al., *RNAalifold: improved consensus structure prediction for RNA alignments*. BMC Bioinformatics, 2008. **9**: p. 474.

[25] K. E. Deigan, et al., *Accurate SHAPE-directed RNA structure determination*. Proc Natl Acad Sci U S A, 2009. **106**(1): p. 97–102.

[26] Z. Sukosd, et al., *Evaluating the accuracy of SHAPE-directed RNA secondary structure predictions*. Nucleic Acids Res, 2013. **41**(5): p. 2807–2816.

[27] X. He, et al., *Inference of RNA structural contacts by direct coupling analysis*. Commun Inf Syst, 2019. **19**(3): p. 279–297.

[28] S. W. Burge, et al., *Rfam 11.0: 10 years of RNA families*. Nucleic Acids Res, 2013. **41**(D): p. D226–D232.

[29] O. Ronneberger, P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Medical Image Computing and Computer-Assisted Intervention, Pt Iii, 2015. **9351**: p. 234–241.

[30] T. Y. Lin, et al., *Focal loss for dense object detection.* IEEE Trans Pattern Anal Mach Intell, 2018.

[31] A. Churkin, L. Weinbrand, and D. Barash, *Free energy minimization to predict RNA secondary structures and computational RNA design.* Methods Mol Biol, 2015. **1269**: p. 3–16.

[32] I. Tinoco, Jr., O.C. Uhlenbeck, and M.D. Levine, *Estimation of secondary structure in ribonucleic acids.* Nature, 1971. **230**(5293): p. 362–367.

[33] B. W. Matthews, *Comparison of the predicted and observed secondary structure of T4 phage lysozyme.* Biochim Biophys Acta, 1975. **405**(2): p. 442–451.

[34] M. Parisien, et al., *New metrics for comparing and assessing discrepancies between RNA 3D structures and models.* RNA, 2009. **15**(10): p. 1875–1885.

[35] H. Yang, et al., *Tools for the automatic identification and classification of RNA base pairs.* Nucleic Acids Res, 2003. **31**(13): p. 3450–3460.

[36] P. Kerpedjiev, S. Hammer, and I.L. Hofacker, *Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams.* Bioinformatics, 2015. **31**(20): p. 3377–3379.

[37] J. Singh, et al., *RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning.* Nat Commun, 2019. **10**: p. 5407.

[38] K. Mao, J. Wang, and Y. Xiao, *Prediction of RNA secondary structure with pseudoknots using coupled deep neural networks.* Biophys Rep, 2020. **6**(4): p. 146–154.

[39] K. Mao, J. Wang, and Y. Xiao, *Length-Dependent Deep Learning Model for RNA Secondary Structure Prediction.* Molecules, 2022. **27**(3): p. 1030.

[40] M. Akiyama, K. Sato, and Y. Sakakibara, *A max-margin training of RNA secondary structure prediction integrated with the thermodynamic model.* J Bioinform Comput Biol, 2018. **16**(6): p. 1840025.

Xiaoling He
School of Physics
Huazhong University of Science and Technology
Wuhan 430074
China
E-mail address: hexl@apm.ac.cn

Xiujuan Ou
School of Physics
Huazhong University of Science and Technology
Wuhan 430074
China
E-mail address: oxj@hust.edu.cn

Hong Yao
School of Physics
Huazhong University of Science and Technology
Wuhan 430074
China
E-mail address: yaohhust@qq.com

Jun Wang
School of Physics
Huazhong University of Science and Technology
Wuhan 430074
China
E-mail address: junwanggwx@gmail.com

Yi Xiao
School of Physics
Huazhong University of Science and Technology
Wuhan 430074
China
E-mail address: yxiao@hust.edu.cn

Received August 16, 2021