OXFORD

DeepRNA-Twist: language-model-guided RNA torsion angle prediction with attention-inception network

Abrar Rahman Abir, Md Toki Tahmid, Rafiqul Islam Rayan, M Saifur Rahman*

Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh

*Corresponding author. Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh.
E-mail: mrahman@cse.buet.ac.bd

Abstract

RNA torsion and pseudo-torsion angles are critical in determining the three-dimensional conformation of RNA molecules, which in turn governs their biological functions. However, current methods are limited by RNA's structural complexity as well as flexibility, with experimental techniques being costly and computational approaches struggling to capture the intricate sequence dependencies needed for accurate predictions. To address these challenges, we introduce DeepRNA-Twist, a novel deep learning framework designed to predict RNA torsion and pseudo-torsion angles directly from sequence. DeepRNA-Twist utilizes RNA language model embeddings, which provides rich, context-aware feature representations of RNA sequences. Additionally, it introduces 2A3IDC module (**A**ttention **A**ugmented Inception Inside Inception with **D**ilated **C**NN), combining inception networks with dilated convolutions and multi-head attention mechanism. The dilated convolutions capture long-range dependencies in the sequence without requiring a large number of parameters, while the multi-head attention mechanism enhances the model's ability to focus on both local and global structural features simultaneously. DeepRNA-Twist was rigorously evaluated on benchmark datasets, including RNA-Puzzles, CASP-RNA, and SPOT-RNA-1D, and demonstrated significant improvements over existing methods, achieving state-of-the-art accuracy. Source code is available at https://github.com/abrarrahmanabir/DeepRNA-Twist

Keywords: RNA language model; torsion and pseudo-torsion angle; inception

Introduction

RNA molecules are fundamental to numerous biological processes and their functionality is closely linked to their three-dimensional structures, akin to proteins. The functionality of RNA critically depends on its structural configuration, highlighting the need for precise characterization of its complex three-dimensional conformation. Traditional experimental methods for determining RNA structure, such as NMR [1], X-ray crystallography [2], and cryo-EM [3], while reliable, are often limited by high costs and time-consuming processes.

Central to RNA's structural complexity are its torsion angles, comprising seven specific angles $(\alpha, \beta, \gamma, \delta, \epsilon, \zeta$ and $\chi)$ along the ribose-phosphate backbone. Additionally, two pseudo-torsion angles $(\eta$ and $\theta)$ provide a simplified represen-tation of the RNA backbone (Fig. 1). These angles are crucial as they dictate RNA's folding patterns and ultimately its three-dimensional shape, enabling RNA to perform diverse biological roles—from catalysis as ribozymes to regulating gene expression through various noncoding RNA mechanisms. Understanding these torsion angles enhances the capability to design therapeutic agents [4], potentially altering their function to treat diseases at a molecular level. Moreover, RNA torsion angles play a critical role in computational biology beyond structure reconstruction. They are widely used in evaluating and comparing predicted RNA 3D models, where they serve as a foundation for similarity metrics like Mean Circular

Quantities (MCQ) and LCS-TA (Longest Common Segment—Torsion Angles) [5]. Their significance extends to fields such as NMR studies, where torsion angles are used to interpret and refine experimental data on RNA conformations.

Predicting RNA torsion angles and structures presents substantial challenges due to RNA's inherent properties and the limitations of current methodologies. Unlike proteins which have only three backbone torsion angles, RNA's structure is determined by a more complex system of torsion angles that contribute to its flexibility and dynamic nature, allowing multiple conformations [6]. This complexity is compounded by RNA's engagement in noncanonical interactions such as Hoogsteen base pairing, base triples, and various loop interactions, vital for its biological functions.

Recent advancements in protein torsion angle prediction have effectively utilized deep learning methods to enhance prediction accuracy [7, 8]. Inspired by this, researchers have employed various deep learning approaches to predict RNA torsion angles as well, albeit this area still remains significantly under-explored. SPOT-RNA-1D [9], the first RNA backbone torsion angle prediction method, utilized a dilated convolutional neural network to predict both torsion and pseudotorsion angles of RNA from single sequence inputs. Currently, the development of language models that can predict RNA structural features solely from sequence data is very limited. RNA-TorsionBERT [10]
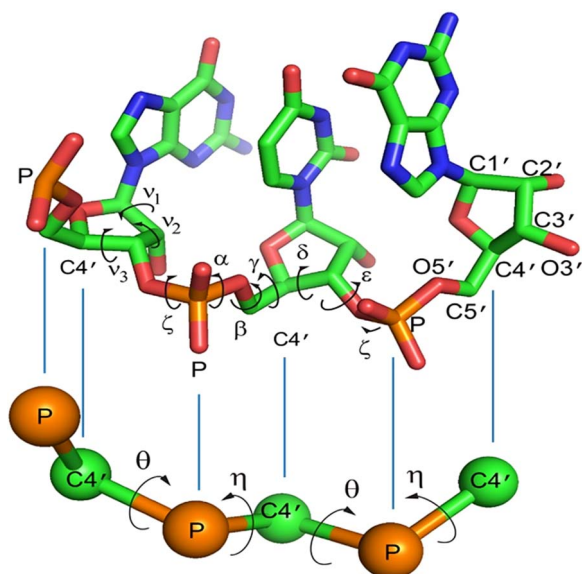
Figure 1. A detailed view of a short RNA sequence (GUG) illustrating all atoms excluding hydrogens along with key atoms and native torsion angles labeled (Top) and pseudo-torsion angles η and θ for the same sequence (Bottom), taken from [11], Copyright 2016 The Authors under the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

model leveraged language model approach to predict torsion and pseudo-torsion angles directly from sequence data. Additionally, this model introduced RNA Torsion-A, a scoring function that assesses the quality of predicted RNA structures using the torsion angles generated by the model, further refining the evaluation of RNA structural predictions.

In this study, we present DeepRNA-Twist, a novel deep learning approach for predicting RNA torsion and pseudo-torsion angles directly from sequence data. A key novelty of DeepRNA-Twist is the utilization of embeddings from an RNA language model (RiNALMo [12]), which has contributed to the model's improved performance. This novel approach to feature representation has not been explored before for RNA torsion angle prediction. Additionally, we have introduced the 2A3IDC (**A**ttention **A**ugmented **I**nception **I**nside **I**nception with **D**ilated **C**NN) module that considerably enhanced the prediction capabilities of DeepRNA-Twist compared with the state-of-the-art models. This module combines inception block and dilated convolution with multi-head attention mechanism. The use of dilated CNN allows the model to capture a broader context and long-range dependencies in the sequence data while enhancing computational efficiency by reducing the number of parameters required to achieve a large receptive field. This enables the model to process sequences more efficiently without sacrificing performance. On the other hand, the multi-head attention mechanism further improves the model's ability to focus on different parts of the sequence simultaneously, enhancing the capture of both local and global structural features. DeepRNA-Twist effectively captures both short- and long-range interactions among nucleotides. We rigorously evaluated DeepRNA-Twist against existing leading approaches on widely recognized benchmark datasets such as RNA-Puzzles [13] and CASP-RNA [14] and SPOT-RNA-1D Dataset [9]. We have employed Mean Absolute Error (MAE) and MCQ as our evaluation metrics. Our results demonstrate that DeepRNA-Twist significantly outperforms existing methods, achieving state-of-the-art accuracy in RNA torsion angle prediction.

Table 1. Training, validation, and test dataset statistics

Dataset	No. of RNA	No. of nucleotide	Max RNA Len.
Training set			
Training [9]	286	21 736	418
Validation set			
VL [9]	30	1215	64
Test sets			
TS1 [9]	63	4725	186
TS2 [9]	30	2040	174
TS3 [9]	54	2187	103
RNA-Puzzle [13]	40	3838	263
CASP-RNA [14]	12	2655	720

Materials and methods
Dataset

To train our model, we have used the training dataset from SPOT-RNA-1D [9]. We then evaluated our model in two independent test datasets: the SPOT-RNA-1D Test dataset and RNA-TorsionBERT [10] Test dataset. Table 1 shows the detailed statistics for the datasets. The maximum length of RNA sequences in Table 1 is important to note because RNA sequence length can impact downstream structural prediction tasks. Longer sequences tend to exhibit increased conformational variability and structural complexity. Including this information thus helps contextualize the diversity of RNA sequences in our dataset and ensures that the model is evaluated across a broad range of sequence lengths. Furthermore, Fig. 2 demonstrates the distribution of all torsion and pseudo-torsion angles of both datasets.

SPOT-RNA-1D dataset

For SPOT-RNA-1D, RNA structures were sourced from the Protein Data Bank (PDB) [15], selecting those with X-ray resolutions finer than 3.5 Å as of 3 October 2020. As described by the authors, these structures were segmented into individual chains using Biopython, then clustered with CD-HIT-EST [16] at an 80% identity threshold to form the training set, with unclustered sequences making up a noncluster set. To refine the data, BLAST-N [17] filtered out sequences with internal or cross-set similarities using an e-value cutoff of 10. The noncluster sequences were subsequently divided into a validation set (VL) and two test sets (TS1 and TS2), ensuring minimal redundancy by constructing covariance models using the INFERNAL tool's cmbuild and cmsearch programs, applying strict e-value cutoffs to eliminate remote homologs. An additional test set (TS3) was later formed from NMR structures, processed similarly to ensure nonredundancy. The final datasets included 286 RNA chains for training, with 30 for validation and 63, 30, and 54 for TS1, TS2, and TS3, respectively, where the maximum sequence length is 418. Native torsion angles were extracted using the DSSR program [18].

RNA-TorsionBERT test dataset

RNA-TorsionBERT [10] employed a combined test set consisting of two prominent datasets: RNA Puzzles [13] and CASP-RNA [14]. This merged dataset consists of 52 structures—40 from RNA Puzzles and 12 from CASP-RNA. The length of sequences of this dataset ranges from 27 to 720 nucleotides.

While there are limited numbers of RNA sequences in all the datasets, the torsion angle prediction is performed on a per
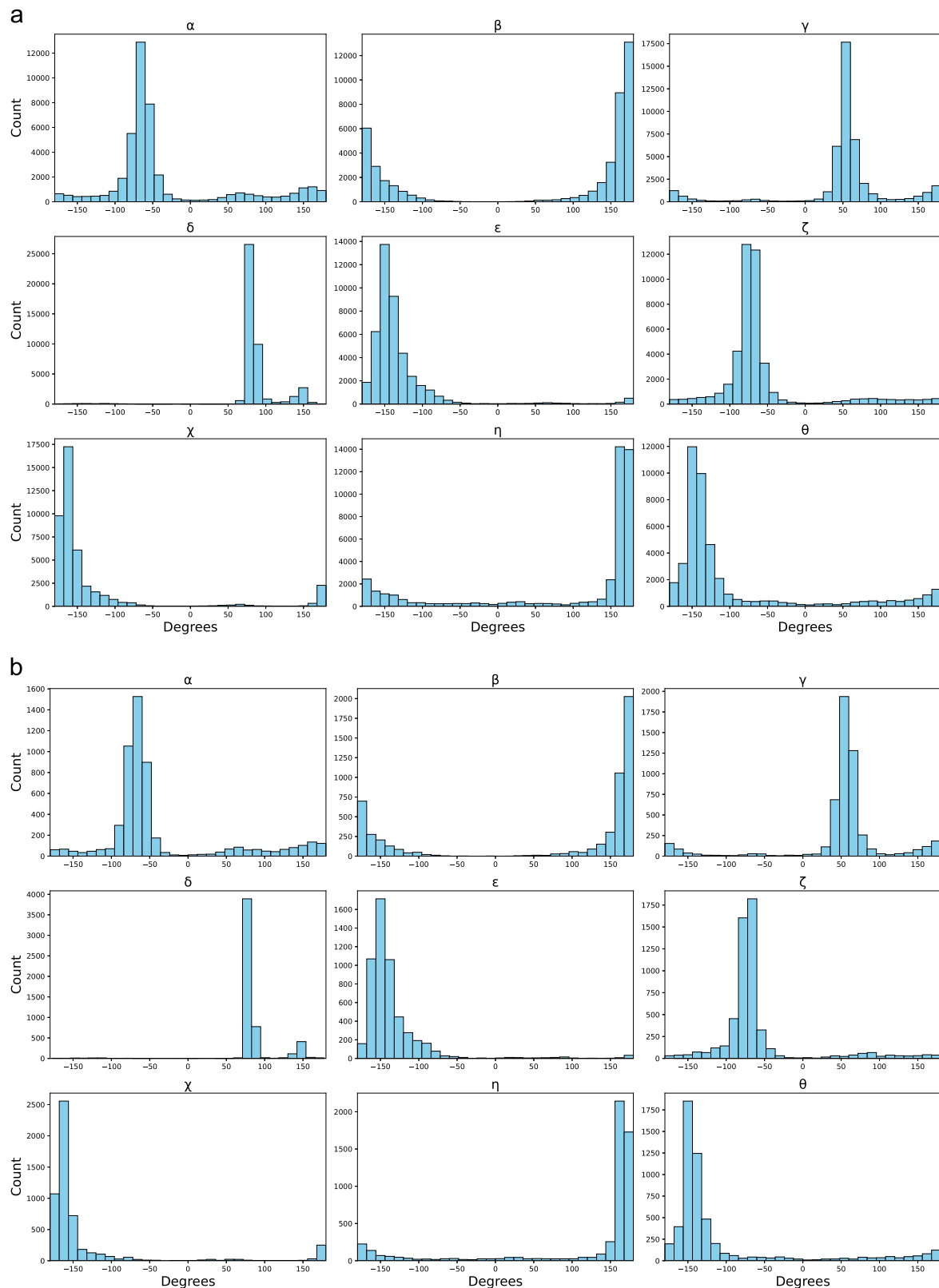
Figure 2. Distribution of all torsion and pseudo-torsion angles for (a) SPOT-RNA-1D dataset and (b) RNA-TorsionBERT dataset.

nucleotide basis. As demonstrated in Table 1, our training set has a sizable amount of nucleotides (21 736). Also, all our test sets combined contain a total of 15 445 nucleotides. Therefore the dataset is well poised for deep learning architecture, and the multiple test sets are able to check the generalizability of our model.

Feature representation

DeepRNA-Twist takes an RNA sequence feature vector $X = \{x_1, x_2, \ldots, x_i, x_{i+1}, \ldots, x_N\}$ as input, where x_i is the vector corresponding to the ith nucleotide of that RNA. Inspired by the success of protein language models in capturing biological patterns and in downstream tasks, we used RiNALMo, an RNA

language model developed in [12], to generate input features. We chose RiNALMo because it is the largest RNA language model developed to date, with 650 million parameters and pretrained on 36 million noncoding RNA sequences sourced from multiple databases. RiNALMo generates a sequence of embedding vectors $X = \{x_1, x_2, \ldots, x_N\}$, where each $x_i \in \mathbb{R}^{d_{\text{RiNALMo}}}$ and $d_{\text{RiNALMo}} = 1280$ represent the features for ith nucleotide of RNA sequence.

Overview of DeepRNA-Twist framework

We propose DeepRNA-Twist, a novel deep learning framework designed to predict RNA torsion angles directly from RNA sequences by effectively capturing both local and global structural patterns within the sequences. As shown in Fig. 4, the pipeline begins by encoding each nucleotide with a feature vector derived from RiNALMo [12], which serves as the foundation for all subsequent computations. These feature vectors are first refined through a Transformer Encoder layer that employs multi-head self-attention and feed-forward networks, thereby capturing essential contextual relationships among nucleotides while maintaining stability via residual connections and layer normalization. The refined representations are then processed through two consecutive 2A3IDC modules, which integrate inception blocks, dilated convolutional layers, and attention mechanism. This combination enables the framework to extract multi-scale features and model long-range dependencies effectively. In these modules, parallel convolutional pathways with different dilation rates capture varying levels of local and global information, while the integrated attention mechanisms further refine these features by focusing on the most relevant positional relationships within the RNA sequence. Following these processing stages, the resulting feature maps are passed through a 1D convolutional layer and an additional attention module before being fed into a dense layer with 18 regression nodes with tanh activation function. These nodes predict the sine and cosine values for each of the nine torsion angles per nucleotide. The training objective is to minimize the mean squared error (MSE) between the predicted and true sine and cosine values across all nucleotides. This comprehensive pipeline enables DeepRNA-Twist to robustly predict RNA torsion angles, providing a strong foundation for downstream structural reconstruction and analysis. Next, we present detailed description of different modules of DeepRNA-Twist.

Transformer encoder layer

DeepRNA-Twist incorporates a Transformer Encoder Layer [19] to process the RNA sequence feature vector $X = \{x_1, x_2, \ldots, x_N\}$, where each x_i is a 1280-dimensional vector corresponding to the ith nucleotide, provided by the RiNALMo [12]. Each Transformer Encoder Layer consists of two main components: a multi-head self-attention mechanism and a position-wise feed-forward network. The multi-head self-attention mechanism refines the representation of each nucleotide by attending to all other nucleotides in the sequence. For each nucleotide x_i, the self-attention mechanism produces an intermediate representation $\mathbf{a}_i$ that captures contextual dependencies. Following the self-attention mechanism, each intermediate representation $\mathbf{a}_i$ is independently passed through a position-wise feed-forward network, resulting in a transformed representation $\mathbf{z}_i$. Formally, this can be expressed as

$$\mathbf{z}_i = \text{FFN}(\mathbf{a}_i),$$

where FFN represents the feed-forward network. Each sub-layer in the encoder, including the self-attention and feed-forward

networks, incorporates residual connections followed by layer normalization. This configuration stabilizes the learning process and enhances the model's ability to capture complex dependencies across the RNA sequence. The overall transformation of x_i through the encoder can be summarized as

$$\mathbf{z}_i = \text{LayerNorm}(\mathbf{a}_i + \text{FFN}(\mathbf{a}_i)),$$

where $\mathbf{a}_i$ is obtained from the multi-head self-attention mechanism applied to x_i. This process results in the refined representation of the initial embeddings, which captures both local and global structural information of the RNA sequence.

Multi-head attention module

The multi-head attention module, denoted by Multi-HeadAttention(.), is designed to dynamically weigh the importance of different elements in the input data, adjusting the focus based on the input's context. This module integrates a positional encoding sub-module, which provides crucial positional information to enhance the model's ability to capture sequential relationships.

Positional encoding: to encode positional information [19], the positional encoding function PE_p for a position p is defined as

$$\text{PE}(p, 2i) = \sin\left(\frac{p}{10\,000^{2i/d_{\text{feature}}}}\right);$$

$$\text{PE}(p, 2i+1) = \cos\left(\frac{p}{10\,000^{2i/d_{\text{feature}}}}\right),$$

where i is the dimension. This allows the model to learn to attend to relative positions. The input $\mathbf{X}$ is augmented with the positional encoding, resulting in a new representation:

$$\mathbf{X}_{\text{pos}} = \mathbf{X} + \text{PE}_{\text{pos}}$$

This new representation $\mathbf{X}_{\text{pos}}$ retains both the information from previous layers and the positional information of each element.

Projection to Query, Key, and Value

The input sequence feature vector $\mathbf{X} = \{x_1, x_2, \ldots, x_N\}$ is first augmented with positional encoding to obtain $\mathbf{X}_{\text{pos}}$. This augmented input is then linearly transformed to create three matrices: Query ($\mathbf{Q}$), Key ($\mathbf{K}$), and Value ($\mathbf{V}$):

$$\mathbf{Q} = \mathbf{X}_{\text{pos}}\mathbf{W}^Q, \quad \mathbf{K} = \mathbf{X}_{\text{pos}}\mathbf{W}^K, \quad \mathbf{V} = \mathbf{X}_{\text{pos}}\mathbf{W}^V,$$

where $\mathbf{W}^Q$, $\mathbf{W}^K$, and $\mathbf{W}^V$ are learnable parameter matrices.

Attention calculation

The multi-head self-attention mechanism allows the model to jointly attend to information from different representation subspaces. The input is split into h heads, and the attention scores are computed for each head using the scaled dot-product of $\mathbf{Q}$ and $\mathbf{K}$, followed by scaling and masking. For the ith head:

$$\mathbf{Q}_i = \mathbf{Q}\mathbf{W}_i^Q, \quad \mathbf{K}_i = \mathbf{K}\mathbf{W}_i^K, \quad \mathbf{V}_i = \mathbf{V}\mathbf{W}_i^V$$

The attention scores ($\mathbf{A}_i$) are computed as

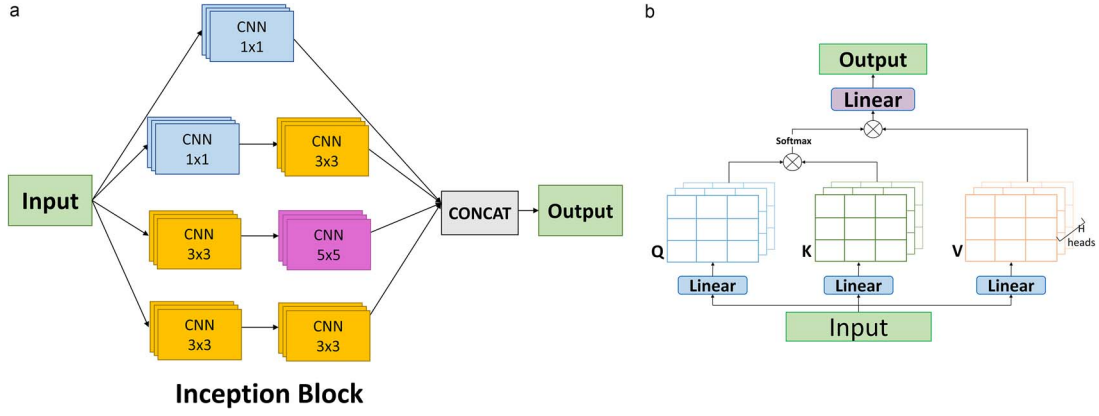$$\mathbf{A}_i = \text{softmax}\left(\frac{\mathbf{Q}_i\mathbf{K}_i^T}{\sqrt{d_k}}\right),$$

Figure 3. (a) Inception Block and (b) Multi-head attention block.

where d_k is the dimension of the keys. The output for each head ($\mathbf{O}_i$) is then computed by

$$\mathbf{O}_i = \mathbf{A}_i \mathbf{V}_i$$

The outputs from all heads are concatenated to form the final output of the multi-head attention mechanism:

$$\mathbf{O}_{\text{concat}} = \text{Concat}(\mathbf{O}_1, \mathbf{O}_2, \ldots, \mathbf{O}_h)$$

The concatenated output is then linearly transformed:

$$\mathbf{O}_{\text{final}} = \mathbf{O}_{\text{concat}} \mathbf{W}^O,$$

where $\mathbf{W}^O$ is a learnable parameter matrix. The resulting output is then subjected to dropout and batch normalization to enhance training stability and performance.

By integrating positional encoding and using multiple attention heads, the attention module can effectively capture both the content and positional relationships within the input, making it suitable for processing RNA sequences.

2A3IDC module

The 2A3IDC module integrates inception blocks, dilated convolutional layers, and attention mechanisms to process input data through two parallel paths, enhancing the model's ability to capture both local and global patterns. This design builds on the 2A3I (Attention Augmented Inception Inside Inception) module that was very successful in protein secondary structure prediction [20]. Each path in the 2A3IDC module begins with an inception block (Fig. 3), consisting of four parallel convolutional pathways, to capture multi-scale features. Following the inception block, a dilated convolutional layer is added. Given an input feature map X:

$$X_{Inception} = Inception(X)$$
$$X_{Dilated} = DilatedConv(X_{Inception}, d)^,$$

where d is the dilation rate, allowing the model to capture long-range dependencies efficiently without increasing the number of parameters. The two parallel paths of the 2A3IDC module use different dilations rates: 2 and 5. The output from the dilated convolutional block is then passed through a multi-head attention module. This mechanism focuses on different parts of the sequence simultaneously, enhancing the model's ability to capture both local and global structural features. Finally, the outputs from both parallel paths ($X_{Attention1}$, $X_{Attention2}$) are concatenated and batch-normalized to form the final representation:

$$X_{Attention} = MultiHeadAttention(X_{Dilated})$$
$$X_{Final} = BatchNorm(Concat(X_{Attention1}, X_{Attention2}))$$

Training details

The loss function used to train DeepRNA-Twist is the MSE between the predicted and true sine and cosine values of the torsion angles across all nucleotides of all the RNAs. Let N be the number of RNAs, L_n the number of nucleotides in the nth RNA in the dataset. Let $A = 9$ be the number of torsion angles per nucleotide. Each angle has both sine and cosine values, leading to $2A = 18$ values per nucleotide. The MSE loss function is defined as

$$\mathcal{L}_{\text{MSE}} = \frac{1}{\sum_{n=1}^{N} L_n \times 2A} \sum_{n=1}^{N} \sum_{l=1}^{L_n} \sum_{a=1}^{2A} \left(\hat{y}_{n,l,a} - y_{n,l,a} \right)^2,$$

where $y_{n,l,a}$ and $\hat{y}_{n,l,a}$ are the true and predicted sine or cosine values corresponding to the torsion angles for the lth nucleotide in the nth RNA. Here, a ranges from 1 to 18, indexing the sine and cosine values for each of the nine torsion angles. We trained DeepRNA-Twist for 120 epochs. We used the Adam optimizer with a learning rate of 0.0001.

Evaluation metric
Mean absolute error

The predictive performance of RNA torsion angle is quantitatively assessed using the MAE. MAE is calculated for each torsion angle across all nucleotides in the sequence, considering the periodicity of the angles. The MAE for a specific torsion angle θ is defined as follows:

$$\text{MAE}(\theta) = \frac{1}{N} \sum_{i=1}^{N} \min \left(\Delta\theta_i, 360° - \Delta\theta_i \right),$$

where $\Delta\theta_i = |\theta_{\text{pred},i} - \theta_{\text{true},i}|$ is the absolute difference between the predicted angle $\theta_{\text{pred},i}$ and the true experimentally determined angle $\theta_{\text{true},i}$ and N is the total number of nucleotides in the RNA sequence for which torsion angles are predicted. Lower MAE indicated better prediction.
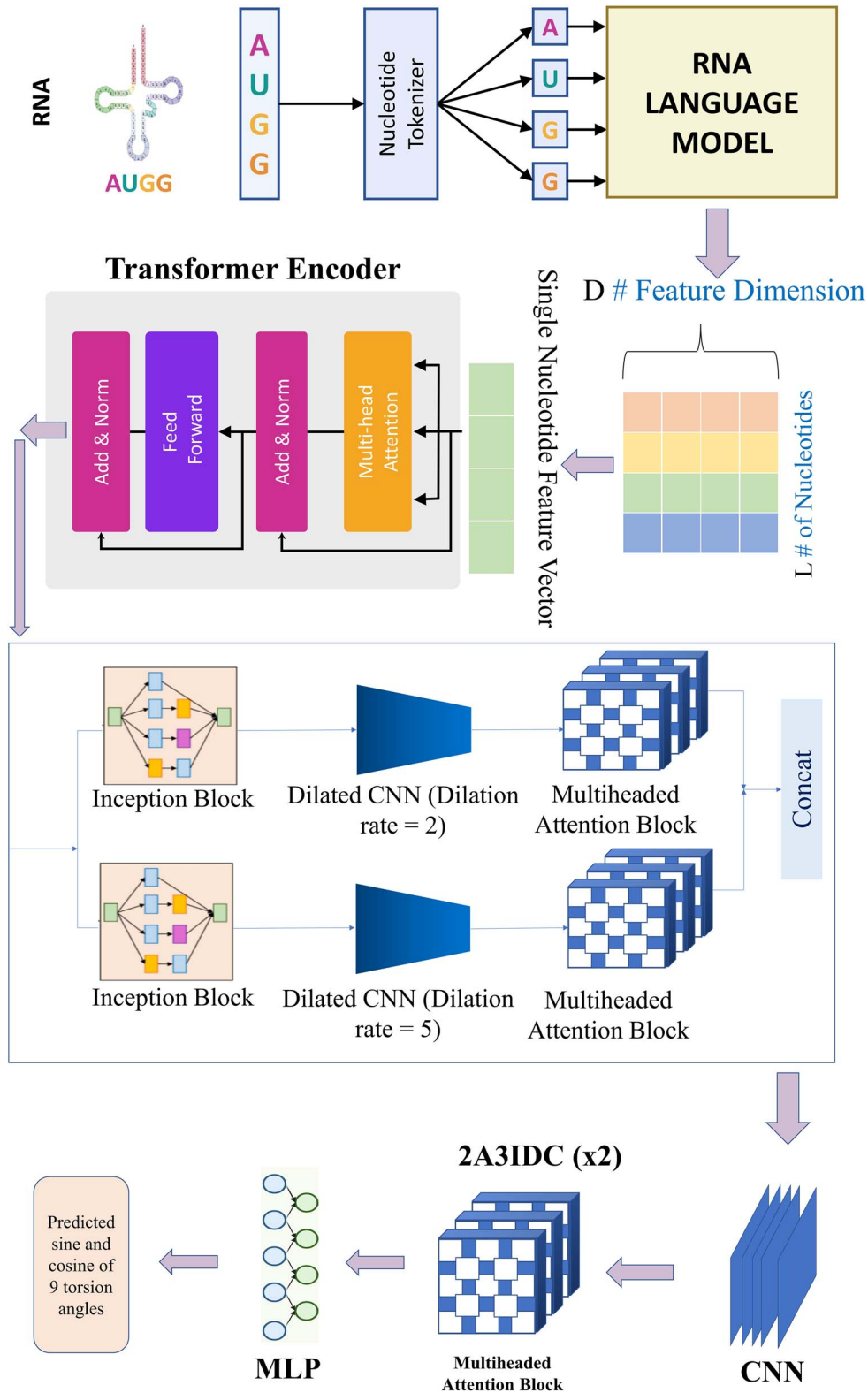
Figure 4. DeepRNA-Twist architecture.

Mean circular quantities

The MCQ is used to measure the angular similarity between actual and predicted RNA structures, S and S', respectively, based on their torsion angles. When the RNA structure consists of n residues, its trigonometric representation can be expressed as a matrix containing $9n$ values of torsion angles t_{ij}, where

$i = 1,\ldots,n,\ j = 1,\ldots,|T|,$ and T is the set of torsion angles defined for the structure. To calculate the MCQ between two structures of equal length, the angular differences $\Delta(t_{ij}, t'_{ij})$ for corresponding torsion angles are computed while respecting the periodicity of angular values. The MCQ metric aggregates these differences using the sine and cosine components as

Table 2. Performance comparison of DeepRNA-Twist and SPOT-RNA-1D according to MAE in degrees on SPOT-RNA-1D test sets. Values of SPOT-RNA-1D are reported from [9]

	Model	α	β	γ	δ	ϵ	ζ	χ	η	θ
VL	DeepRNA-Twist	40.32	17.40	29.76	14.83	17.12	34.68	20.20	29.47	35.93
	SPOT-RNA-1D	45.18	20.58	33.88	17.99	20.72	37.50	23.01	33.55	37.02
TS1	DeepRNA-Twist	39.16	17.67	28.77	12.79	16.98	29.50	16.33	27.12	28.52
	SPOT-RNA-1D	43.94	21.94	32.98	14.61	20.69	33.27	19.59	30.25	32.91
TS2	DeepRNA-Twist	35.71	16.12	24.29	13.90	14.44	25.00	15.56	25.19	24.96
	SPOT-RNA-1D	39.50	18.92	29.47	16.01	17.46	28.91	18.20	28.14	30.25
TS3	DeepRNA-Twist	34.13	16.89	29.54	11.30	19.32	24.17	13.92	22.20	25.84
	SPOT-RNA-1D	37.89	21.04	34.68	13.83	22.32	27.87	17.01	25.31	27.22

follows:

$$\text{MCQ}(S, S') = \arctan\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{|T|} \sin \Delta(t_{ij}, t'_{ij})}{\sum_{i=1}^{n}\sum_{j=1}^{|T|} \cos \Delta(t_{ij}, t'_{ij})}\right),$$

where $\Delta(t_{ij}, t'_{ij})$ is the minimum angular difference defined as

$$\Delta(t_{ij}, t'_{ij}) = \min\left(|t_{ij} - t'_{ij}|, 360° - |t_{ij} - t'_{ij}|\right).$$

Earlier works, such as SPOT-RNA-1D, only used the MAE metric. We on the other hand have used MCQ in addition to MAE to assess the proposed model and make a comparative analysis with the state-of-the-art models. This makes our analysis more rigorous.

Results

We evaluated our model on test datasets of SPOT-RNA-1D and RNA-TorsionBERT, the results of which are reported below.

Performance on SPOT-RNA-1D dataset

DeepRNA-Twist was trained on train dataset of SPOT-RNA-1D [9] and validated using validation dataset VL. Table 2 shows the performance comparison between DeepRNA-Twist and SPOT-RNA-1D on VL and three independent test sets TS1, TS2, and TS3. In the VL, our model demonstrated a noticeable improvement over SPOT-RNA-1D, showing percentage reductions in MAE that ranged from ~10% to 15% across most torsion angles. Particularly, the reductions for the angles α, η, and γ were among the most significant. This trend of enhancement is almost consistent across the test sets, indicating robustness in handling both standard and pseudo-torsion angles. TS1 displayed similarly substantial improvements with reductions particularly pronounced for the α and θ angles. In TS2, our model excelled particularly in the prediction of γ and θ, achieving more than a 15% improvement in MAE compared with SPOT-RNA-1D. TS3, derived from structurally diverse NMR datasets, showed our model's adaptability with significant reductions in MAE for β and γ. The performance in this set validates our model's utility in handling RNA datasets generated from different experimental techniques.

The improvements in MAE were statistically significant, as evidenced by P-values from paired t-tests (Table 3). Moreover, the prediction performance is not sensitive to the length of RNA sequences, as shown in Fig. 5. The MAE remains relatively stable across different RNA length bins, with no significant upward or downward trends. This indicates that DeepRNA-Twist generalizes well across RNA molecules of varying lengths.
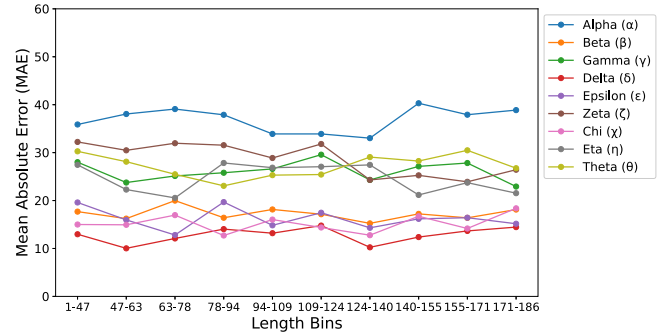

Figure 5. Variation of MAE with respect to the length (number of nucleotides) of RNA sequences.

Our analysis confirms that torsion angles with narrower distributions, such as δ, ϵ, and χ, tend to be more predictable. As shown in Fig. 2, these angles exhibit sharp peaks and narrowly distributed values, indicating lower variability across RNA structures. This observation aligns with SPOT-RNA-1D [9], which suggests that angles with more constrained distributions are easier to predict. Conversely, angles like α, ζ, and θ display broader distributions with widespread values, making them inherently more challenging to predict. Figure 2 illustrates how these angles exhibit greater variability, often spanning a wide range of values. However, as shown in Table 2, our model achieves lower MAE for these more variable angles compared with SPOT-RNA-1D, demonstrating its robustness in handling torsion angles with higher degrees of flexibility.

Performance for nucleotides in various pairing interactions

The performance of our model was evaluated on torsion angle prediction across different RNA interactions within the test datasets TS1, TS2, and TS3. Each dataset was analyzed based on the nucleotides involved in various structural interactions: unpaired bases, lone pairs, pseudoknots, multiplets, noncanonical pairs, and canonical nested base pairs. The predictive accuracy of our model was compared with the SPOT-RNA-1D as shown in Tables 4, 5, and 6.

Analysis revealed that nucleotides involved in tertiary interactions, such as lone pairs, pseudoknots, multiplets, and noncanonical pairs, consistently presented greater challenges in torsion angle prediction across all three test sets. This was evident from the consistently higher MAEs for these interactions compared with canonical nested base pairs, which had the lowest MAEs, indicating their relative predictability due to less structural complexity. In TS1, the unpaired bases showed the most significant

Table 3. P-values from one-tailed paired t-tests comparing RNA torsion angle predictions by DeepRNA-TWIST with SPOT-RNA-1D and RNA-TorsionBERT

Model	α	β	γ	δ	ϵ
SPOT-RNA-1D	0.00038	0.00108	0.00026	0.00185	0.00019
RNA-TorsionBERT	0.0000871	0.0000005311	0.0000974	0.00536	0.00000585

Model	ζ	χ	η	θ
SPOT-RNA-1D	0.00037	0.00011	0.00050	0.03203
RNA-TorsionBERT	0.0000405	0.0004884	0.0000708	0.00000136

Table 4. Performance comparison of DeepRNA-Twist and SPOT-RNA-1D according to MAE in degrees based on various pairing interactions on SPOT-RNA-1D test set—TS1. Values of SPOT-RNA-1D are reported from [9]

Angles	Model	Unpaired	Lone pairs	Pseudoknots	Multiplets	Noncanonical pairs	Canonical nested pairs
α	DeepRNA-Twist	59.55	51.26	51.16	46.92	44.24	32.02
	SPOT-RNA-1D	62.05	54.26	54.16	49.92	47.24	35.02
β	DeepRNA-Twist	25.93	24.96	25.32	24.52	21.14	15.32
	SPOT-RNA-1D	28.43	27.96	28.32	27.52	24.14	18.32
γ	DeepRNA-Twist	36.26	37.11	35.42	37.73	32.08	26.17
	SPOT-RNA-1D	39.76	40.61	38.42	40.73	35.08	29.17
δ	DeepRNA-Twist	17.53	17.93	19.03	14.97	14.45	8.59
	SPOT-RNA-1D	20.53	20.93	22.03	16.97	16.45	10.59
ϵ	DeepRNA-Twist	25.67	25.82	22.76	22.52	20.42	13.46
	SPOT-RNA-1D	29.17	28.32	25.76	24.52	22.42	15.46
ζ	DeepRNA-Twist	49.10	47.14	43.95	33.84	35.44	17.63
	SPOT-RNA-1D	54.10	51.64	47.95	36.84	38.44	19.63
χ	DeepRNA-Twist	28.33	24.18	18.91	19.19	20.77	11.43
	SPOT-RNA-1D	30.83	27.18	21.91	22.19	22.77	13.43
η	DeepRNA-Twist	51.08	39.26	37.54	32.84	32.87	16.24
	SPOT-RNA-1D	54.58	42.76	40.54	34.84	34.87	18.24
θ	DeepRNA-Twist	51.20	40.52	39.15	31.80	36.83	18.02
	SPOT-RNA-1D	56.70	44.52	42.15	33.80	39.83	20.02

Table 5. Performance comparison of DeepRNA-Twist and SPOT-RNA-1D according to MAE in degrees based on various pairing interactions on SPOT-RNA-1D test set—TS2. Values of SPOT-RNA-1D are reported from [9]

Angles	Model	Unpaired	Lone pairs	Pseudoknots	Multiplets	Noncanonical pairs	Canonical nested pairs
α	DeepRNA-Twist	57.26	50.62	54.40	47.43	41.32	25.03
	SPOT-RNA-1D	60.26	53.62	57.40	50.43	44.32	28.03
β	DeepRNA-Twist	21.49	23.17	26.72	20.69	17.78	13.14
	SPOT-RNA-1D	24.49	26.17	28.72	22.69	19.78	15.14
γ	DeepRNA-Twist	35.89	36.49	32.46	34.66	29.99	21.91
	SPOT-RNA-1D	37.89	39.49	35.46	36.66	31.99	23.91
δ	DeepRNA-Twist	19.74	20.89	20.77	18.13	17.08	9.43
	SPOT-RNA-1D	22.74	23.89	23.77	21.13	20.08	11.43
ϵ	DeepRNA-Twist	24.92	21.20	16.43	16.20	17.43	11.30
	SPOT-RNA-1D	26.92	24.20	18.43	18.20	19.43	12.30
ζ	DeepRNA-Twist	45.34	47.20	30.74	38.04	40.14	13.73
	SPOT-RNA-1D	48.34	50.20	32.74	41.04	43.14	14.73
χ	DeepRNA-Twist	25.67	25.00	18.06	21.32	21.74	10.77
	SPOT-RNA-1D	27.67	28.00	21.06	24.32	23.74	12.77
η	DeepRNA-Twist	49.92	36.04	39.38	31.45	30.03	14.01
	SPOT-RNA-1D	52.92	39.04	42.38	33.45	32.03	16.01
θ	DeepRNA-Twist	49.40	41.06	32.54	31.03	34.05	15.21
	SPOT-RNA-1D	52.40	44.06	35.54	34.03	37.05	16.21

challenge in prediction due to more variability and flexibility in structure, with MAEs of 59.55 for α, considerably higher than those for canonical nested pairs, which had an MAE of 32.02 for the

same angle. This pattern was consistent across all angles and other test datasets. TS3, characterized by different experimental derivation techniques (NMR vs. X-ray crystallography for TS1 and

Table 6. Performance comparison of DeepRNA-Twist and SPOT-RNA-1D according to MAE in degrees based on various pairing interactions on SPOT-RNA-1D test set—TS3. Values of SPOT-RNA-1D are reported from [9]

Angles	Model	Unpaired	Lone pairs	Pseudoknots	Multiplets	Noncanonical pairs	Canonical nested pairs
α	DeepRNA-Twist	64.82	48.32	49.06	43.16	49.68	22.44
	SPOT-RNA-1D	68.82	51.32	52.06	46.16	52.68	25.44
β	DeepRNA-Twist	31.74	29.53	29.42	27.95	26.72	13.40
	SPOT-RNA-1D	33.74	31.53	31.42	29.95	28.72	15.40
γ	DeepRNA-Twist	54.96	47.84	30.42	43.50	41.05	22.61
	SPOT-RNA-1D	58.96	50.84	32.42	46.50	44.05	24.61
δ	DeepRNA-Twist	24.39	15.84	16.67	13.22	13.20	8.75
	SPOT-RNA-1D	27.39	17.84	18.67	16.22	15.20	10.75
ϵ	DeepRNA-Twist	30.83	25.01	19.40	22.73	21.94	17.68
	SPOT-RNA-1D	32.83	27.01	21.40	24.73	22.94	18.68
ζ	DeepRNA-Twist	58.55	37.07	27.20	33.15	33.95	14.99
	SPOT-RNA-1D	61.55	40.07	29.20	35.15	34.95	15.99
χ	DeepRNA-Twist	25.66	21.59	13.42	18.40	19.01	11.77
	SPOT-RNA-1D	27.66	23.59	15.42	20.40	21.01	12.77
η	DeepRNA-Twist	51.43	28.11	30.02	26.60	25.14	14.01
	SPOT-RNA-1D	54.43	30.11	32.02	28.60	27.14	16.01
θ	DeepRNA-Twist	55.84	34.86	29.55	27.70	25.10	15.21
	SPOT-RNA-1D	58.84	36.86	31.55	29.70	27.10	16.21

TS2), still maintained the general trend where more complex tertiary interactions resulted in higher prediction errors.

Comparative analysis between our model and SPOT-RNA-1D [9] demonstrated that despite the inherent challenges posed by complex interactions, our model achieved consistently lower MAEs across most interactions and torsion angles, indicating an overall improvement in prediction accuracy. Our findings suggest that the complexity of nucleotide interactions significantly influences the difficulty of predicting torsion angles in RNA structures. More specifically, interactions that introduce greater structural complexity and variability tend to result in higher prediction errors. The comprehensive evaluation across multiple test sets confirms the robustness of our model, particularly in its superior performance over the baseline in handling the intricate and variable interactions within RNA structures. We conducted this experiment on RNA-TorsionBERT dataset as well and noticed the same trend (Table 7).

Performance on RNA-TorsionBERT dataset

Table 8 presents the performance of our method, DeepRNA-Twist, in predicting RNA torsion angles compared with various state-of-the-art approaches including RNA-TorsionBERT [10], SPOT-RNA-1D [9], and methods from the recent State-of-the-RNArt [21] benchmark on RNA-TorsionBERT dataset. DeepRNA-Twist demonstrates superior accuracy with the lowest mean MAE of 19.78° across all torsion angles. This represents a significant improvement over RNA-TorsionBERT [10], which records a mean MAE of 22.6°, and a more pronounced advantage over SPOT-RNA-1D (mean MAE of 30.0°). Specifically, DeepRNA-Twist achieves the best performance in predicting δ with an MAE of 12.1°, surpassing RNA-TorsionBERT's MAE of 13.6° and greatly outperforming SPOT-RNA-1D's 28.3°. The largest improvement over RNA-TorsionBERT [10] is observed in the γ angle, where DeepRNA-Twist reduces the MAE by nearly 4°. Moreover, across other angles such as χ, ϵ, and β, DeepRNA-Twist consistently maintains lower MAEs compared with all benchmarked methods. We also notice the same trend in the relationship between prediction difficulty and the distribution of angles. Angles associated with wider distribution, such as α,

γ, and ζ, are more difficult to predict compared with angles with narrower distribution such as δ and ϵ. DeepRNA-Twist consistently shows statistically significant improvements in MAE across all torsion angles in this dataset over all other methods (Table 3).

Performance comparison based on MCQ

Table 9 shows the MCQ of DeepRNA-Twist, SPOT-RNA-1D, and RNA-TorsionBERT on both datasets. The MCQ metric quantifies the dissimilarity between the actual and predicted three-dimensional structures. A higher MCQ value indicates greater dissimilarity between the actual and predicted structures, while a lower MCQ value signifies greater similarity. On both datasets, DeepRNA-Twist achieves the least MCQ, which highlights the effectiveness of our approach in predicting RNA three-dimensional structures.

Ablation study

To evaluate the contributions of different components within the DeepRNA-Twist architecture, we conducted an ablation study in two parts: (1) a comparison between RiNALMo embeddings and one-hot encoded features, and (2) an assessment of individual architectural components by systematically removing them. The results are summarized in Tables 10 and 11, which include results for both RNA-TorsionBERT and SPOT-RNA-1D datasets. We first compared the performance of DeepRNA-Twist model using RiNALMo embeddings with a simpler feature encoding approach with one-hot encoded nucleotides. As shown in Table 10, the model using RiNALMo embeddings consistently outperforms the one-hot encoded model on both RNA-TorsionBERT and SPOT-RNA-1D datasets. Furthermore, we also investigated the performance of different RNA language models as shown in Fig. 6. For this experiment, we chose BiRNA-BERT [31] and RNA-FM [32] as the two competing language models. RiNALMo turns out to be the best-performing model on both datasets and thus it validates our design choice.

Table 7. Performance comparison of DeepRNA-Twist and RNA-TorsionBERT according to MAE in degrees based on various pairing interactions on RNA-TorsionBERT test set. Values of RNA-TorsionBERT are obtained by running RNA-TorsionBER

Angles	Model	Unpaired	Lone pairs	Pseudoknots	Multiplets	Noncanonical pairs	Canonical nested pairs
α	DeepRNA-Twist	62.03	46.55	55.78	42.28	49.43	30.02
	RNA-TorsionBERT	66.16	57.11	50.97	51.70	52.46	34.34
β	DeepRNA-Twist	24.09	21.88	30.93	30.07	24.93	15.93
	RNA-TorsionBERT	29.31	28.79	30.44	26.04	28.58	19.68
γ	DeepRNA-Twist	32.55	38.87	41.37	43.03	26.55	21.72
	RNA-TorsionBERT	42.77	44.82	40.40	43.60	38.77	29.31
δ	DeepRNA-Twist	23.52	13.27	15.25	9.14	11.74	6.66
	RNA-TorsionBERT	25.81	20.62	26.92	13.53	15.30	14.60
ϵ	DeepRNA-Twist	20.02	26.37	27.70	22.15	20.37	15.14
	RNA-TorsionBERT	26.88	29.81	21.20	29.12	20.77	16.72
ζ	DeepRNA-Twist	52.17	41.70	43.74	32.15	39.00	21.00
	RNA-TorsionBERT	52.40	48.01	53.93	37.25	34.74	15.88
χ	DeepRNA-Twist	32.09	20.54	18.94	16.60	16.67	12.52
	RNA-TorsionBERT	31.44	29.60	18.66	20.94	26.01	19.00
η	DeepRNA-Twist	55.18	39.53	34.95	37.60	36.89	18.11
	RNA-TorsionBERT	52.72	41.64	46.53	39.40	31.48	17.68
θ	DeepRNA-Twist	54.71	38.11	41.08	35.91	32.58	16.54
	RNA-TorsionBERT	60.09	49.40	45.88	35.86	45.61	20.81

Table 8. Performance comparison according to MAE in degrees on RNA-TorsionBERT test sets

Model	α	β	γ	δ	ϵ	ζ	χ	η	θ	Mean
DeepRNA-Twist	**34.7**	**16.8**	**25.5**	**12.1**	**13.4**	**23.4**	**12.6**	**17.2**	**22.3**	**19.78**
RNA-TorsionBERT* [10]	37.3	19.6	29.4	13.6	16.6	26.6	14.7	20.1	25.4	22.6
SPOT-RNA-ID* [9]	50.7	30.1	35.8	28.3	21.9	29.8	20.9	25.0	27.5	30.0
lsRNA1* [22]	53.1	26.2	39.2	14.6	25.9	38.4	19.0	27.6	31.7	30.6
AlphaFold* 3 [23]	51.2	26.5	38.0	17.0	21.3	38.3	20.5	30.2	36.3	31.1
RNAJP* [24]	56.8	31.6	41.9	18.6	27.7	42.1	21.1	29.8	33.3	33.7
Vfold-Pipeline* [25]	55.4	28.1	38.8	20.1	32.0	43.1	25.0	34.2	41.5	35.4
RNAComposer* [26]	58.9	31.5	48.2	17.9	27.1	40.0	24.8	32.8	38.6	35.5
MC-Sym* [27]	72.2	31.2	61.5	31.8	25.6	49.4	29.6	41.2	44.6	43.0
3dRNA* [28]	65.7	40.0	53.6	33.7	40.3	52.7	27.2	40.2	45.6	44.3
trRosettaRNA* [29]	67.6	36.8	67.0	26.4	32.9	54.3	58.3	38.2	48.3	47.8
Rhofold* [30]	94.3	66.0	76.2	60.6	52.8	72.6	32.4	42.0	46.6	60.4

*Reported from the preprint version of [10], available at https://www.biorxiv.org/content/10.1101/2024.06.06.597803v1

Table 9. Performance comparison based on MCQ

(a) RNA-TorsionBERT Dataset

Model	MCQ
DeepRNA-Twist	**15.13**
SPOT-RNA-1D	19.40
RNA-TorsionBERT	17.40

(b) SPOT-RNA-1D Dataset

Model	MCQ
DeepRNA-Twist	**17.93**
SPOT-RNA-1D	24.47
RNA-TorsionBERT	22.72

Table 10. Feature encoding ablation: mean MAE of nine torsion angles

(a) RNA-TorsionBERT Dataset

Feature Encoding	Mean MAE (°)
RiNALMo	**19.78**
One-Hot Encoding	23.45

(b) SPOT-RNA-1D Dataset

Feature Encoding	Mean MAE (°)
RiNALMo	**23.58**
One-Hot Encoding	27.31

In the second part of our ablation study, we evaluated the contributions of key components within DeepRNA-Twist architecture by systematically removing them. Each component's removal allowed us to assess its impact on the overall model performance, as reflected by the mean MAE of all torsion angles. Table 11 presents the results of the component removal experiments. The removal of the 2A3IDC modules led to the highest increase in mean MAE across both datasets, emphasizing their critical role in capturing both local and global dependencies within the sequence. Removing the Transformer Encoder Layer

Table 11. Component ablation: mean MAE of nine torsion angles

(a) RNA-TorsionBERT Dataset

Model Variant	Mean MAE (°)
Full Model	**19.78**
Without Transformer Encoder	22.90
Without 2A3IDC Modules	24.31
Without Multi-head Attention	21.86

(b) SPOT-RNA-1D Dataset

Model Variant	Mean MAE (°)
Full Model	**23.58**
Without Transformer Encoder	26.34
Without 2A3IDC Modules	28.72
Without Multi-head Attention	25.91

Table 12. Comparison among 2A3IDC, 2A3I, and 3I modules

(a) RNA-TorsionBERT Dataset

Module	Mean MAE (°)	Total parameters
2A3IDC	**19.78**	2371 820
2A3I	26.91	3153 020
3I	28.72	1707 620

(b) SPOT-RNA-1D Dataset

Module	Mean MAE (°)	Total arameters
2A3IDC	**23.58**	2371 820
2A3I	27.84	3153 020
3I	29.10	1707 620
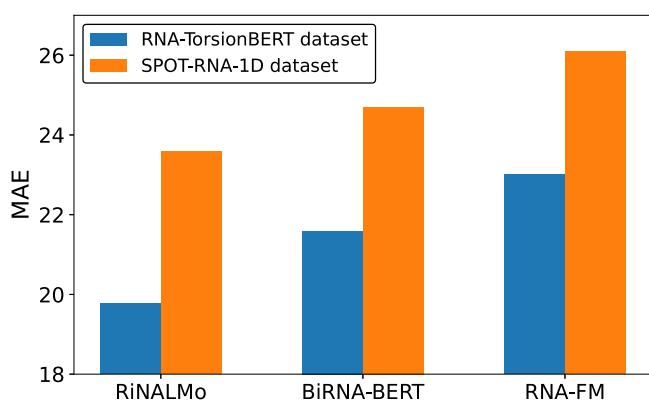
Figure 6. MAE for different RNA Language Models.

and Multi-head Attention also resulted in performance drops, highlighting their importance in the model's architecture.

To further investigate the effectiveness of 2A3IDC module, we replaced this module with similar type of architectures (2A3I [20] and 3I [33]) while keeping all the other parts of the network intact. Table 12 records the results, clearly showing the superiority of 2A3IDC module over the other two. In particular, its superior performance over 2A3I is commendable since it has far less parameters, demonstrating its efficiency. This is attributed to its design, which combines Inception networks, dilated convolutions, and multi-head attention. The dilated CNNs capture broader context and long-range dependencies with fewer parameters, while the multi-head attention enhances the capture of local and global features. In contrast, 3I, which lacks an attention mechanism, has the fewest parameters but delivers the worst performance. These results show 2A3IDC's ability to balance computational efficiency and accuracy. The ablation study confirms that the 2A3IDC modules and language model embeddings are vital, validating our design choices in DeepRNA-Twist.

Case study

In this case study, we started with native RNA structures (as PDB files) with PDB IDs 4R4V and 7PTK and obtained predicted backbone torsion angles using DeepRNA-Twist. These predicted angles were then applied to the corresponding nucleotides of native structures in PyMOL [34] using the *set_dihedral* function, generating new RNA conformations. The reconstructed structures

were subsequently superimposed onto their native PDB structures (see Fig. 7), and the root mean square deviation (RMSD) values were calculated to quantify the accuracy of the predictions. All modifications, visualization, and evaluation were performed within PyMOL. In the superimposed visualizations, the native structures are displayed in cyan, while the predicted structures are shown in magenta, allowing for a direct comparison of the model's predictions with the experimental data. For 7PTK, the RMSD between the predicted and native structures was 6.59 Å, and for 4R4V, the RMSD was 3.31 Å. We also employed two other torsion angle prediction models—SPOT-RNA-1D and RNA-TorsionBERT and two recent sequence to 3D structure prediction models—AlphaFold3 [23] and RhoFold+ [35] for this experiment as shown in Fig. 8. For the sequence to 3D structure prediction models, the RNA sequence was given as input and the predicted 3D structure was the output, which we used for comparison. These results demonstrate the capability to closely align predicted RNA structures with their experimentally determined counterparts, highlighting the effectiveness of DeepRNA-Twist in RNA torsion angle prediction over the state-of-the-art models.

Discussion

In this study, we introduced DeepRNA-Twist, a novel deep learning approach for predicting RNA torsion and pseudo-torsion angles solely from sequence data. DeepRNA-Twist integrates several key modules, each contributing unique strengths. The transformer encoder captures complex dependencies and contextual relationships, essential for RNA torsion angle prediction. The 2A3IDC modules extract multi-scale features and recognize long-range dependencies while refining focus on critical structural variations through multi-head attention mechanisms. Attention modules throughout the architecture dynamically weigh contributions, ensuring significant features are prioritized. Our evaluations on benchmark datasets such as RNA-Puzzles [13], CASP-RNA [14], and SPOT-RNA-1D [9] demonstrate that DeepRNA-Twist significantly outperforms existing methods, achieving state-of-the-art accuracy. On the SPOT-RNA-1D dataset, the model showed 10% to 15% MAE improvements, depicting its robustness and accuracy. Angles with wider distribution like α, ζ, and θ showed notable improvements, whereas angles like δ and ϵ, which have narrower distribution, were easier to predict. Despite complexities in predicting angles for nucleotides involved in tertiary interactions, DeepRNA-Twist consistently achieved lower MAEs compared with other methods. Further analysis on RNA-Puzzles and CASP-RNA
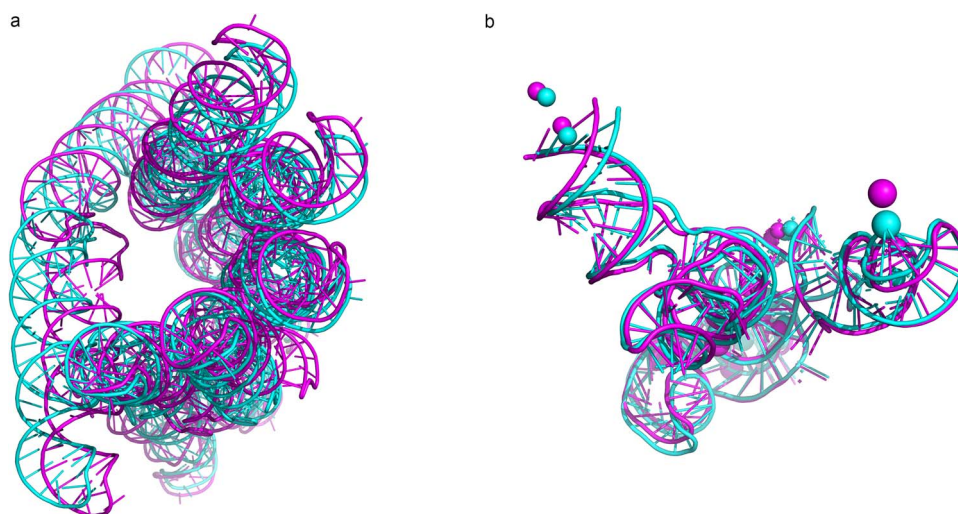
Figure 7. Visualization of the superposition of RNA structures: (a) 7PTK and (b) 4R4V using the torsion angles estimated by DeepRNA-Twist (magenta color), compared with the structures using the native angles obtained from PDB (cyan color).
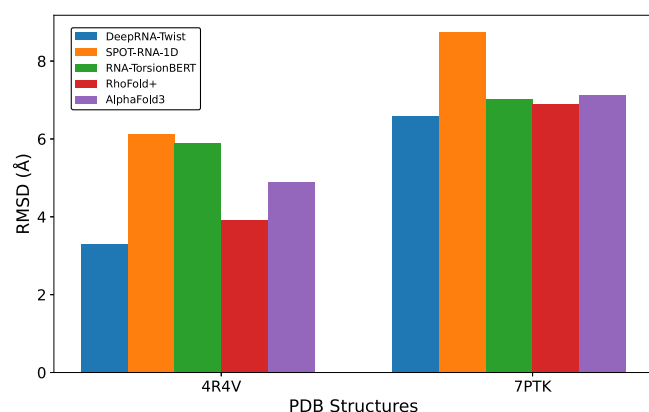

Figure 8. Comparison of RMSD for 4R4V and 7PTK.

datasets highlights that DeepRNA-Twist outperforms deep learning based state-of-the-art methods, including RNA-TorsionBERT [10], SPOT-RNA-1D [9], and AlphaFold 3 [23], with an overall MAE reduction of more than 10° and a 35% improvement in accuracy compared with template-based and ab initio methods. Moreover, we show the effectiveness of DeepRNA-Twist in predicting RNA three-dimensional structures by using MCQ as evaluation metric. These results highlight the model's superior performance and generalizability. Despite these advancements, inherent limitations remain. Reducing MAE below a certain threshold is challenging due to RNA structural complexity and variability. Angles like α and γ with wider distribution are particularly difficult to predict accurately, necessitating continued research, the integration of additional biophysical constraints, and more comprehensive datasets. In our future work, we aim to address these issues to further improve prediction accuracy.

Key Points

- We present DeepRNA-Twist, a novel deep learning framework for predicting RNA torsion and pseudo-torsion angles from RNA sequence.

- We introduce 2A3IDC module (**A**ttention **A**ugmented **I**nception **I**nside **I**nception with **D**ilated **C**NN), combining inception networks, dilated CNNs, and multi-head attention to capture both short- and long-range dependencies in RNA sequences efficiently.
- DeepRNA-Twist achieves state-of-the-art accuracy in RNA torsion angle prediction, outperforming existing methods on benchmark datasets like RNA-Puzzles, CASP-RNA, and SPOT-RNA-1D.
- We provide interpretations of DeepRNA-Twist's performance based on angle distribution and nucleotide interactions, demonstrating significant improvements, especially for angles with broader distributions and complex tertiary interactions which are difficult to predict.
- DeepRNA-Twist is freely available as an easy-to-use script at https://github.com/abrarrahmanabir/DeepRNA-Twist.

Funding

References

1. Scott LG, Hennig M. RNA structure determination by NMR. *Methods Mol Biol* 2008;**452**:29–61.

2. Jackson RW, Smathers CM, Robart AR. General strategies for RNA X-ray crystallography. *Molecules* 2023;**28**:2111. https://doi.org/10.3390/molecules28052111

3. Zhang K, Li S, Kappel K. *et al.* Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nat Commun* 2019;**10**:5511.

4. Kim YK. RNA therapy: rich history, various applications and unlimited future prospects. *Exp Mol Med* 2022;**54**:455–65. https://doi.org/10.1038/s12276-022-00757-5

5. Mackowiak M, Adamczyk B, Szachniuk M. *et al.* RNAtango: analysing and comparing RNA 3D structures via torsional angles. *PLoS Comput Biol* 2024;**20**:e1012500. https://doi.org/10.1371/journal.pcbi.1012500

6. Frellsen J, Moltke I, Thiim M. *et al.* A probabilistic model of RNA conformational space. *PLoS Comput Biol* 2009;**5**:e1000406. https://doi.org/10.1371/journal.pcbi.1000406

7. Xu YC, ShangGuan TJ, Ding XM. *et al.* Accurate prediction of protein torsion angles using evolutionary signatures and recurrent neural network. *Sci Rep* 2021;**11**:21033. https://doi.org/10.1038/s41598-021-00477-2

8. Hanson J, Paliwal K, Litfin T. *et al.* Improving prediction of protein secondary structure, backbone angles, solvent accessibility and contact numbers by using predicted contact maps and an ensemble of recurrent and residual convolutional neural networks. *Bioinformatics* 2019;**35**:2403–10. https://doi.org/10.1093/bioinformatics/bty1006

9. Singh J, Paliwal K, Singh J. *et al.* RNA backbone torsion and pseudotorsion angle prediction using dilated convolutional neural networks. *J Chem Inf Model* 2021;**61**:2610–22. https://doi.org/10.1021/acs.jcim.1c00153

10. Bernard C, Postic G, Ghannay S. *et al.* RNA-TorsionBERT: leveraging language models for RNA 3D torsion angles prediction. *Bioinformatics* 2025;**41**:btaf004.

11. Dawson WK, Maciejczyk M, Jankowska EJ. *et al.* Coarse-grained modeling of RNA 3D structure. *Methods* 2016;**103**:138–56. https://doi.org/10.1016/j.ymeth.2016.04.026

12. Penić RJ, Vlašić T, Huber RG, Wan Y, Šikić M. Rinalmo: General-purpose rna language models can generalize well on structure prediction tasks. arXiv preprint arXiv:2403.00043. 2024.

13. Magnus M, Antczak M, Zok T. *et al.* RNA-Puzzles toolkit: a computational resource of RNA 3D structure benchmark datasets, structure manipulation, and evaluation tools. *Nucleic Acids Res* 2020;**48**:576–88. https://doi.org/10.1093/nar/gkz1108

14. Das R, Kretsch RC, Simpkin AJ. *et al.* Assessment of three-dimensional RNA structure prediction in CASP15. *Proteins* 2023;**91**:1747–70. https://doi.org/10.1002/prot.26602

15. Rose PW, Prlić A, Altunkaya A. *et al.* The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res* 2016;**45**:gkw1000.

16. Fu L, Niu B, Zhu Z. *et al.* CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 2012;**28**:3150–2. https://doi.org/10.1093/bioinformatics/bts565

17. Altschul SF, Gish W, Miller W. *et al.* Basic local alignment search tool. *J Mol Biol* 1990;**215**:403–10. https://doi.org/10.1016/S0022-2836(05)80360-2

18. Lu X-J, Bussemaker HJ, Olson WK. DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res* 2015;**43**:gkv716. https://doi.org/10.1093/nar/gkv716

19. Vaswani A, Shazeer N, Parmar N. *et al.* Attention is all you need. *Advances in neural information processing systems* 2017;**30**.

20. Uddin MR, Mahbub S, Rahman MS. *et al.* SAINT: self-attention augmented inception-inside-inception network improves protein secondary structure prediction. *Bioinformatics* 2020;**36**:4599–608. https://doi.org/10.1093/bioinformatics/btaa531

21. Bernard C, Postic G, Ghannay S. *et al.* State-of-the-RNArt: benchmarking current methods for RNA 3D structure prediction. *NAR Genomics Bioinf* 2024;**6**:lqae048. https://doi.org/10.1093/nargab/lqae048

22. Zhang D, Li J, Chen S-J. IsRNA1: de novo prediction and blind screening of RNA 3D structures. *J Chem Theory Comput* 2021;**17**:1842–57. https://doi.org/10.1021/acs.jctc.0c01148

23. Abramson J, Adler J, Dunger J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493–500. https://doi.org/10.1038/s41586-024-07487-w

24. Li J, Chen S-J. RNAJP: enhanced RNA 3D structure predictions with noncanonical interactions and global topology sampling. *Nucleic Acids Res* 2023;**51**:3341–56. https://doi.org/10.1093/nar/gkad122

25. Li J, Zhang S, Zhang D. *et al.* Vfold-Pipeline: a web server for RNA 3D structure prediction from sequences. *Bioinformatics* 2022;**38**:4042–3. https://doi.org/10.1093/bioinformatics/btac426

26. Popenda M, Szachniuk M, Antczak M. *et al.* Automated 3D structure composition for large RNAs. *Nucleic Acids Res* 2012;**40**:e112–2. https://doi.org/10.1093/nar/gks339

27. Parisien M, Major F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* 2008;**452**:51–5. https://doi.org/10.1038/nature06684

28. Wang J, Wang J, Huang Y. *et al.* 3dRNA v2.0: an updated web server for RNA 3D structure prediction. *Int J Mol Sci* 2019;**20**:4116. https://doi.org/10.3390/ijms20174116

29. Wang W, Feng C, Han R. *et al.* trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nat Commun* 2023;**14**:7266. https://doi.org/10.1038/s41467-023-42528-4

30. Shen T, Hu Z, Peng Z. *et al.* E2Efold-3D: End-to-End Deep Learning Method for Accurate de Novo RNA 3D Structure Prediction arXiv preprint arXiv:2207.01586. 2022.

31. Tahmid MT, Shahgir HS, Mahbub S. *et al.* BiRNA-BERT allows efficient RNA language modeling with adaptive tokenization. *bioRxiv* 2024;2024–07.

32. Chen J, Hu Z, Sun S. *et al.* Interpretable RNA foundation model from unannotated data for highly accurate RNA structure and function predictions arXiv preprint arXiv:2204.00300. 2022.

33. Fang C, Shang Y, Xu D. MUFold-SS: new deep inception-inside-inception networks for protein secondary structure prediction. *Proteins* 2018;**86**:592–8. https://doi.org/10.1002/prot.25487

34. Schrodinger LLC. The pymol molecular graphics system *Version*, 1. 2015;8.

35. Shen T, Zhihang H, Sun S. *et al.* Accurate RNA 3D structure prediction using a language model-based deep learning approach. *Nat Methods* 2024;1–12.