IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 21 May 2024, accepted 10 June 2024, date of publication 13 June 2024, date of current version 26 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3413818

RESEARCH ARTICLE

A Comprehensive but Effective Share-Dependent Bidirectional Framework for Relational Triple Extraction

YOUREN CHEN[1,2], YONG LI[1], AND MING WEN[2]
[1]College of Computer Science and Technology, Xinjiang Normal University, Xinjiang 830054, China
[2]Xinjiang Electronics Research Institute, Xinjiang 830011, China

Corresponding author: Yong Li (liyong@live.com)

This work was supported in part by Xinjiang Key Research and Development Program under Grant 2022B01007-3, in part by the Natural Science Foundation of Xinjiang Uygur Autonomous Region under Grant 2022D01A225, and in part by the National Natural Science Foundation of China under Grant 62241209.

ABSTRACT In the evolving landscape of information extraction systems, relational triple extraction has emerged as a pivotal component, especially with the rise of tagging-based methods. This research addresses the inherent limitations of traditional unidirectional and subject-centric models, which often struggle with complex relational structures, leading to compromised accuracy. We introduce a bidirectional extraction framework, a significant departure from conventional models. This novel approach diminishes the heavy dependence on subject extraction, thereby enhancing both the accuracy and robustness of the process. At the heart of our method is the parallel identification of subject-object pairs, underpinned by a shared encoder that adeptly merges feature fusion to augment process interdependence. Our model incorporates a relation dual-stream extraction mechanism, integrating a coordinate attention system essential for assigning complex relationships to each entity pair. This is further refined through a dependency encoder, ensuring a nuanced and precise extraction process. The innovation of this approach lies in its dual-stream framework, meticulously designed to handle the nuances of shared and dependency structures in relational data. This strategy successfully addresses the challenges of cross-dependencies and the oversight of inter-dependency elements, common in earlier models. Through extensive evaluations across multiple benchmark datasets, this model demonstrates superior performance compared to existing methodologies. Its versatility suggests a significant potential for advancing other tagging-based relational triple extraction methods. Consequently, this research not only establishes a new benchmark in relational triple extraction but also indicates a promising direction for the development of more sophisticated and accurate information extraction systems. This has broad implications for the future of natural language processing and knowledge graph construction.

INDEX TERMS Relational triple extraction, joint extraction of entities and relations, deep learning.

I. INTRODUCTION

Relational Triple Extraction (RTE), a crucial task in natural language processing, aims to distill structured knowledge in the form of (subject, relation, object) triples from unstructured text. These triples, where the subject and object are entities connected by a semantic relation, encapsulate factual information [1]. Relation Triple Extraction is pivotal in transforming unstructured natural language texts into structured (subject, relation, object) tuples, essential for applications like automatic knowledge graph construction. Early RTE approaches, using pipeline-based models, have shown promising flexibility and efficiency [2]. These methods, integrating named entity recognition and relation classification, efficiently identify entities in text and predict relations for entity pairs. Their design inherently leverages the strengths

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo.

of each component, making them less susceptible to error propagation compared to more complex models. For instance, the triple (Paris, capital of, France) conveys the fact that Washington is the capital of the United States. The significance of RTE has grown in recent years, largely due to its vital role in applications such as the automatic construction of knowledge graphs. This burgeoning interest has spurred the development of various innovative RTE methodologies [3].

Joint extraction methods have evolved to extract entities and relations simultaneously in an end-to-end manner. Among these, label-based joint extraction methods excel in handling complex sentence structures, including overlapping and multiple triples. However, they encounter limitations in ground entity extraction failure, significantly impacting RTE performance [4].

Early RTE Methodologies: The initial phase of RTE research, as referenced in studies [5] primarily utilized a pipeline-based framework. This approach first identified all entities within the text and then predicted relations for each pair of entities. Such methods leveraged the advancements in named entity recognition and relation classification. However, they were hindered by two major shortcomings [6], [7], [8], [9], [10]. Firstly, they failed to acknowledge the interdependence between entity recognition and relation prediction [11], [12], [13], [14], [15]. Secondly, they were significantly affected by error propagation, where inaccuracies in the early stages of the pipeline adversely impacted the subsequent stages [16], [17], [18], [19], [20].

To address these deficiencies, the focus shifted towards joint extraction methods, as seen in [21], [22], [23], and [24]. These methods strive to extract entities and relations concurrently in an end-to-end manner. Such an integrated approach has shown considerable promise, particularly in handling complex sentence structures.

Tagging-Based Joint Extraction Methods: Within the sphere of joint extraction, tagging-based methods [1], [3], [5], [25], [26], [27] have gained prominence due to their impressive performance and capability to extract triples from complex sentences. These complex sentences often contain overlapping triples, where a single entity or an entity pair is involved in multiple relational triples within the same sentence [28], or sentences that encompass multiple triples. Existing tagging-based methods generally decompose the RTE task into two subtasks: one for recognizing all subjects and another for simultaneously recognizing all objects and relations. This is often referred to as the unidirectional extraction framework.

In summary, while RTE methodologies have evolved from pipeline-based to more integrated joint extraction approaches, challenges persist, particularly in the accurate extraction of ground entities [29]. Addressing these challenges is critical for advancing RTE capabilities, thereby enhancing the accuracy and reliability of knowledge extraction from unstructured text. The ongoing development of RTE methodologies reflects a continuous effort to balance efficiency, accuracy, and complexity, essential for the diverse applications that rely on this technology [30]. Despite their successes, these methods are not without limitations. A significant challenge they face is what can be termed as ground entity extraction failure. This issue arises when the failure to accurately extract a subject leads to the subsequent failure in extracting all triples associated with that subject. Table 1 illustrates these scenarios, where triples share one or two entities or are more complex in a sentence. In this context, a ground entity is defined as an entity that is extracted first in a triple. The inability to accurately extract ground entities severely impacts the overall RTE performance [31].

TABLE 1. Examples of normal, EntityPairOverlap(EPO), SingleEntityOverlap(SEO), and HeadTailOverlap(HTO) overlapping patterns.

Type	Sentence	Triple
Normal	[Peking Duck] is a famous dish from [Beijing].	(Peking Duck, Originated in, Beijing)
EPO	[Paris] is the capital city of [France].	(Paris, Capital of, France) (Paris, city of, France)
SEO	[Albert Einstein] was born in [Ulm], a city in [Germany].	(Albert Einstein, Place of birth, Ulm) (Ulm,city of, Germany)
HTO	[Leo Messi] is a renowned football player.	(Leo Messi, Family name, Messi)

In response, BiSDRTE presents a bidirectional shared dependency extraction framework. It initiates by extracting subject-object pairs in both directions, followed by relation prediction. This is augmented by a shared dependency encoder, harmonizing feature extraction and reinforcing the mutual validation of triples. The framework's strength lies in its complementary dual-direction extraction, enhancing overall triple accuracy.

Central to BiSDRTE is the use of a lightweight, yet powerful, coordinate attention affine model. This model adeptly assigns relations to subject-object pairs, harnessing deep interactions without the complexity and resource demands of larger models. The implementation of Residual Networks (RN) within this structure focuses on preserving both original and learned features, boosting the model's learning capacity [32]. Additionally, a weight-adjustable shared perception learning mechanism is introduced to address convergence issues, optimizing learning rates across different modules. Existing research demonstrates that with a lightweight model BERT [8], it is possible to achieve high-quality RTE results without escalating computational resource consumption [1], [2], [4], [24], [25].

II. RELATION WORK

In the field of RTE, our research situates itself within a rapidly evolving landscape, aiming to transform unstructured natural language texts into structured (subject, relation, object) tuples. This transformation is not just a technical challenge but a crucial step in applications like automatic knowledge graph construction. The journey of RTE methodologies from early pipeline-based models to contemporary advanced frameworks reflects a paradigm shift in handling linguistic data, emphasizing efficiency, accuracy, and computational resource management [33].

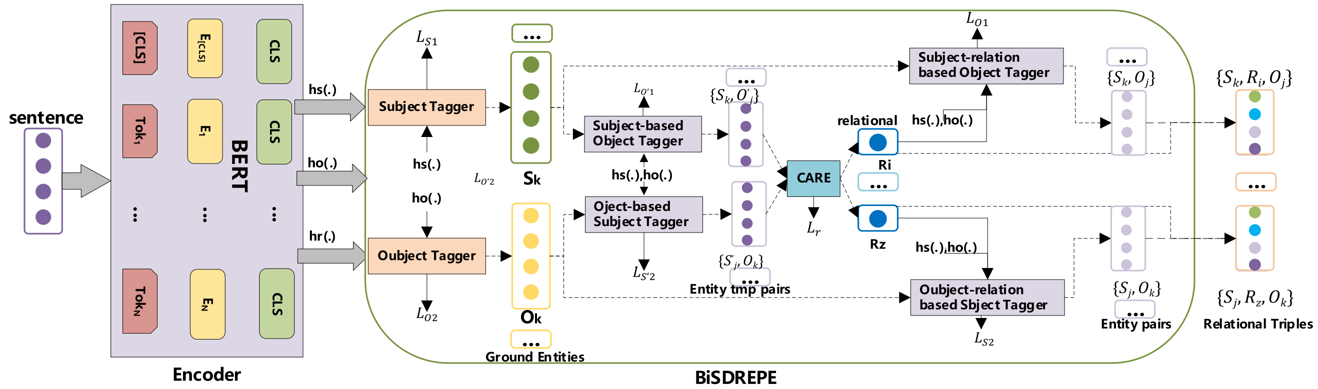

FIGURE 1. BiSDRTE framework, the training process is represented by solid lines, while the inference phase involves dashed lines, indicating the flow and interaction of information during these phases. BiSDRRET, a key component of BiSDRTE, operates under a shared-dependency architecture, encompassing a bidirectional and quadrilateral framework.

The inception of RTE was marked by pipeline-based models, which integrated named entity recognition and relation classification in a sequential manner. These models were lauded for their flexibility and efficiency, adept at identifying entities in texts and predicting relationships. One of their key strengths was the modular design, which minimized error propagation - a significant issue in more complex models. Yet, these early approaches lacked the sophistication to handle intricate relational structures and context dependencies adequately [34], [35], [36].

Subsequently, RTE saw the emergence of joint extraction methods, revolutionizing the field by extracting entities and relations simultaneously in an end-to-end manner. Among these, label-based joint extraction methods stood out for their capability to handle complex sentence structures, including scenarios with overlapping and multiple triples. However, these methods grappled with limitations, particularly in accurately extracting ground entities, which significantly impacted RTE performance [26].

Label-Based Approaches: Pioneering studies such as [4] transformed RTE into a labeling task. This approach enabled the direct extraction of entities and their relationships. Recent advancements, like CasRel [27] is a model based on a cascading binary relation extraction framework, designed to efficiently handle relation extraction tasks by simultaneously identifying entities and the relationships between them in text, but often encountered dependency propagation errors and challenges in refining entity predictions.

Table Filling Approaches: Illustrated by studies like [11], [14], [23], and [33] this methodology maps relationships onto a matrix, converting RTE into a table completion task. Despite its innovative nature, its computational inefficiency has limited its application, particularly in specialized industrial settings.

Sequence-to-Sequence (Seq2Seq) Model-Based Approaches: Studies such as [19] and [28] employed Seq2Seq models, viewing triples as token sequences and redefining RTE as a generative task. However, the reliance on extensive pre-trained models imposed significant computational resource demands.

Recent research has also delved into alternative RTE methods. Notable among these is the unified framework for extracting explicit and implicit relationship triples [4], [7], a three-dimensional perspective on RTE [22], and a focus on latent relationships and global correspondences [37]. These studies have broadened the RTE landscape, introducing new dimensions and perspectives in understanding and extracting relational data.

Our paper primarily discusses label-based methods suitable for small-domain industrial RTE, proposing an innovative framework. We emphasize the efficacy of streamlined models like BERT, which achieve high performance without the computational overhead of more extensive models. This consideration is especially pertinent in RTE applications where resource efficiency is a primary concern.

In conclusion, our research contributes to the field of RTE by introducing BiSDRTE, a framework that not only addresses the limitations of existing methodologies but also sets a new benchmark in terms of efficiency and effectiveness. This work, therefore, stands as a significant advancement in the ongoing development of RTE technologies, particularly suited for applications demanding high accuracy and computational efficiency.

III. BiSDRTE MODEL

The architecture of BiDRRTE, depicted in Figure 1, encompasses three principal components: a BERT-based Encoder component, a Bidirectional Entity Pair Extraction component (abbreviated as BiDREPE), and a Coordinate Attention-based Relation Extraction component (abbreviated as CARE). During the training phase, the modules within BiDREPE employ multitask learning and feature interaction, while also incorporating the concept of residual networks to preserve the original data features. This approach provides the advantage of allowing each module to be trained with actual input data, thereby facilitating the development of a more reliable model. However, during the inference phase, BiDREPE and CARE operate sequentially.

Subject-to-Object (s2o) and Object-to-Subject (o2s) processes complement the BiSDEPE model by directly identifying relationships between subjects and objects without the initial recognition of predicates, streamlining the extraction of simpler relational patterns. Subject-Relation-Predicate-to-Object (sp2o) and Object-Relation-to-Subject (op2s): These directions enhance the model's capability by integrating the recognition of predicates. In sp2o, BiSDREPE identifies subjects and predicates first, followed by extracting related objects. Conversely, in op2s, the model starts with objects and predicates, subsequently extracting related subjects. This approach emphasizes the importance of understanding the relationships between objects and their respective subjects, mediated by predicates. The shared-dependency architecture is particularly evident in these directions, where predicates act as a bridge connecting subjects and objects, facilitating more accurate interpretation and extraction by integrating all four directions (s2o, o2s, sp2o, op2s), BiSDEPE offers an efficient and comprehensive method for identifying and analyzing complex relationships within the text. The shared encoder component not only enhances the model's efficiency but also promotes synergy among the different extraction directions. This integration results in a powerful tool capable of capturing intricate entity relationships, making BiSDEPE a significant advancement in processing and understanding complex textual data. To provide a detailed explanation of the model structure, we divided the model into seven core components based on Figure 1: A.Encoder, B.Subject Tagger, C.Subject-based Object Tagger, D.Expanded Analysis of the Subject-Relation-Based Object Tagger, E.CARE: Coordinate Attention Mechanism in BiSDRTE, and F.Share-aware Learning Mechanism.

A. ENCODER

In this study, we employ a pre-trained BERT (Cased) [8] model as the cornerstone for encoding tokens within an input sentence, producing an initial representation for each token, symbolized as $h(.) \in \Re^{d_h}$, the superiority of BERT is also proven in literature [4]. These sequences serve as specialized context features tailored to the unique attributes of subjects, objects, and relations. This strategic deviation from the norm is predicated on our understanding that each element within a triple possesses distinct characteristics. Unlike conventional approaches adopted by leading methods such as TPLinker [27], CasRel [25], PRGC [36], which rely on a unified feature set for subjects, objects, and relations, our methodology distinguishes itself by generating three separate token representation sequences.

At its foundation, the model leverages advanced techniques for token representation, denoted as $h_{token_S}, h_{token_O}, h_{token_r}$, utilizing a calculated approach. To derive the initial representation of a token, $h_{token_X}^{original}$, the model employs BERT, a powerful language representation model. The computation within BERT can be abstracted as follows:

$$h_{token_X}^{BERT} = \mathrm{BERT}\left(X_{input}\right) \tag{1}$$

Here, X_{input} represents the input tokens fed into BERT, and $h_{token_X}^{BERT}$ is the output contextual embedding for $token_X$. BERT captures deep contextual relations among tokens through a series of self-attention and feedforward neural network layers, producing embeddings that are rich in semantic and syntactic information.

Following the extraction of $h_{token_X}^{BERT}$, the representation is further refined using a linear transformation, specified by the equation:

$$h_{token_X} = W_{token_X} \cdot h_{token_X}^{BERT} + b_{token_X} \tag{2}$$

In this equation, $W_{(.)} \in \Re^{d_h \times d_h}$ is a trainable matrix, $b_{(.)} \in \Re^{d_h}$ is a bias vector, and d_h is the dimensionality of the feature space. This linear transformation tailors the BERT-generated embeddings to the specific needs of the BiSDRTE model, ensuring that each token's representation is optimized for subsequent processing.

Considering the interdependence between entities in a triple, the model further refines these representations. In BERT, the CLS(Classification) token is a special symbol added to the beginning of each input sequence, used primarily for classification tasks. The object's CLS vector, h_{cls}, is added to the subject's representation (h) to enhance contextual awareness, a process similarly applied to the object representation for symmetry and balance. This dual enhancement strategy underscores the model's innovative approach to leveraging deep linguistic contexts and entity relationships.

B. SUBJECT TAGGER

The model features a key component known as the Subject Tagger. This module employs a binary tagging approach to extract all subject entities from a given input sentence. For each token within the sentence, the Subject Tagger assigns two distinct probabilities: one indicating the likelihood of the token being the starting point of a subject entity, and the other denoting the probability of it being the endpoint. These probabilities are calculated using the following equation:

$$\begin{aligned} p_s^{i,\,start} &= \sigma\left(W_s^{start} h_s^i + b_s^{start}\right) \\ p_s^{i,\,end} &= \sigma\left(W_s^{end} h_s^i + b_s^{end}\right) \end{aligned} \tag{3}$$

where $p_s^{i,start}$ and $p_s^{i,end}$ represent the probabilities of the i-th to-ken being the start token and end token of a subject respectively. $W_s^{(\cdot)} \in R^{1 \times d_h}$ is a trainable matrix, $b_s^{(\cdot)} \in R^1$ is a bias vector. All equations of this paper, σ denotes a sigmoid activation function.

In this study, we use a simple 1/0 tagging scheme, which means a token will be assigned a tag if its probability exceeds a certain threshold or a 0 tag otherwise.

In this model, a straightforward binary tagging scheme is adopted: a token is tagged as 1 (indicating the start or end of a subject) if its calculated probability surpasses a predefined threshold, and as 0 otherwise.

C. SUBJECT-BASED OBJECT TAGGER

The BiSDEPE model includes a Subject-based Object Tagger. This component is designed to identify all object entities concerning previously extracted subjects, following an iterative tagging process. For each selected subject, every token in the input sentence is evaluated, assigning two probabilities to determine if it represents the start or end of an object related to that subject. These probabilities are computed similarly to the Subject Tagger, adhering to the shared-dependency architecture of the model, method as shown in equation (4).

$$\begin{aligned} v_s^{s-k} &= \text{maxpool}\left(h_s^{s-k\_tart}, \cdots, h_s^{s-k-\text{ end}}\right) \\ p_{o'}^{\text{istart}} &= \sigma\left(W_{o'}^{\text{start}}\left(h_{o'}^{i} \circ V_s^{s-k}\right) + b_{o'}^{\text{start}}\right) \\ p_{o'}^{\text{i, end}} &= \sigma\left(W_{o'}^{\text{end}}\left(h_{o'}^{i} \circ V_s^{s-k}\right) + b_{o'}^{\text{end}}\right) \end{aligned} \tag{4}$$

where $h_s^{s\_k\_start}, \ldots, h_s^{s\_-k\_end}$ are the vector representations of the tokens in the k-th subject, so $v_s^{s\_-k}$ can be viewed as a representation for the k-th subject. The maxpool means the max-pooling operation. p_o^{istart} and $p_o^{i,end}$ are the probabilities of the i-th token being the start and end tokens of an object related to the k-th subject respectively. O denotes a hadamard product operation. $W_o^{(.)} \in R^{1\times d_h}$ is a trainable matrix, and $b_o^{(.)} \in R^1$ is a bias vector.

Additionally, BiSDEPE's architecture and its components reflect a sophisticated approach to entity pair extraction, leveraging a bidirectional framework and a shared encoder to efficiently process and analyze textual data.

Understood, expand on these critical components of the BiSDEPE model with a more comprehensive analysis, particularly focusing on the subject-relation-based object Tagger and the nuanced application of cross-entropy-based losses in its learning mechanism.

D. EXPANDED ANALYSIS OF THE SUBJECT-RELATION-BASED OBJECT TAGGER

The subject-relation-based object Tagger in the BiSDEPE model represents a significant advancement in understanding complex text relationships. It enhances the capabilities of the subject-based object Tagger by integrating a deeper understanding of the relational dynamics between subjects and objects:

1) THE TAGGER CALCULATES THE PROBABILITIES OF TOKENS

The tagger calculates the probabilities of tokens being the start or end of an object, incorporating intricate relational context. This sophistication in probability calculation is further enhanced by the model's capability to refine its functions through dependency relationships during the training process. These adjustments allow the model to adapt more accurately to the complexities of natural language, ensuring that its predictions are not only based on surface-level analysis but also deeper syntactic and semantic relationships. The formulae are:

$$\begin{aligned} &p_{o\_start} \\ &= \sigma\left(w_{start} \times \left(\left(h_o \circ v_S^{s-k}\right) + \left(h_o \circ v_S^{s-k} \circ v_p^*\right)\right) + b_{start}\right) \\ &p_{s\_end} \\ &= \sigma\left(w_{end} \times \left(\left(h_o \circ v_S^{s-k}\right) + \left(h_o \circ v_S^{s\text{-}k} \circ v_p^*\right)\right) + b_{end}\right) \end{aligned} \tag{5}$$

In these equations, σ represents the sigmoid function, essential for normalizing the output probabilities to a range between 0 and 1. The vectors h_o are the object vectors, encapsulating the features of the potential object entities. $w_{start}, w_{end} \in R^{1\times d}$ are trainable weight matrices, crucial for the model's adaptability and learning. The bias vectors $b_{start}, b_{end} \in R^1$ are key for adjusting the activation thresholds. The vectors v_k and v_{-k} denote the positive and negative subject-relation vectors, introducing a nuanced understanding of the relational context. The addition of v^*, an extra vector component, further deepens the relational analysis, allowing the Tagger to discern complex interplays between subjects and objects within the text.

2) RELATIONAL CONTEXT ENHANCEMENT

This advanced tagging approach enables the model to not just identify objects but also understand their contextual relationship with the subjects, considering factors like proximity, semantic relatedness, and syntactic dependencies. This understanding is vital for tasks like information extraction, question answering, and relationship extraction in text, where the context and nuances of relationships significantly impact the interpretation and usability of the extracted information.

Cross Entropy-based Losses as mentioned above, all the extraction modules in two directions work in a multi-task learning manner. Thus, both extraction modules in each direction have their loss functions. We denote the losses of the above two tagger modules as $\mathcal{L}_{s1}$ and $\mathcal{L}_{01}$ respectively, and both of them are defined with binary cross entropy-based loss, as shown in equation (6).

$$\begin{aligned} \text{cew}(p,t) &= -\left[\log p + (1-t)\log(1-p)\right] \\ L_o &= \frac{1}{2\times 1}\sum_{m\in\{start,end\}}\sum_{i=1}^{1}\text{cew}\left(p_o^{i,m}, t_o^{i,m}\right) \end{aligned} \tag{6}$$

where cew(p,t) is a binary cross-entropy loss, $p \in (0, 1)$ is the predicted probability and t is the true tag, and 1 is the number of tokens in an input sentence, cew(p,t) is a key component in the model's learning mechanism.

It fine-tunes the model's predictions by penalizing deviations from the actual tags. This loss function is particularly effective for binary classifications, like determining the start or end points of entities, as it ensures that the predicted probabilities align closely with the true binary tags. Similarly, there are two tagger losses in the s, o, s2o, o2s, sp2o, and op2s direction. We denote them as $\mathcal{L}_{s1}$, $\mathcal{L}_{O2}$, $\mathcal{L}_{O'1}$ $\mathcal{L}_{S2}$, $\mathcal{L}_{O1}$, and $\mathcal{L}_{S2}$ respectively and they are computed with a similar method as shown in equation (6).

This loss function adeptly handles the trade-off between precision and recall, reducing the likelihood of false positives (erroneously identifying non-entity tokens as part of an entity) and false negatives (missing actual entity tokens). This balance is crucial for maintaining a high level of accuracy in entity extraction, particularly in complex texts with nuanced entity relationships.

Adaptability and Versatility: Given its foundational role in the model's learning process, the cross-entropy loss function is adaptable to various linguistic contexts and datasets. This versatility is essential for models intended for diverse applications in natural language processing, ensuring consistent performance across different text types and domains.

In summary, the subject-relation-based object Tagger's advanced relational analysis capabilities, combined with the strategic application of cross-entropy-based losses, significantly enhance the BiSDEPE model's ability to accurately and effectively extract and interpret complex entity relationships in textual data.

E. CARE: COORDINATE ATTENTION MECHANISM IN BiSDRTE

The proposed framework makes BiSDEPE output more s-o pairs, where there are many noise pairs. This is harmful to the precision of BiSDRTE. Thus, CARE should have a strong classification ability. Here we use a biaffine and CA(Coordinate Attention) model for the RE(Relation Extraction) module [3], [37]. It maintains a parameter matrix for each relation, and an entity pair will be computed with each relation-specific matrix to determine whether it possesses the corresponding relation or not. Specifically, for an entity pair $\left(s_k, o_j'\right)$, we first obtain the representation vectors $v_r^{s\text{-}k}$ and $v_r^{0'-j}$ its two entities. Then the possibility denoted as P_r^i of $\left(s_k, o_j'\right)$ possessing the i-th relation is computed.

The process is shown in equation (7), where $W_R^1 \in R^{(d_h+1)\times(d_h+1)}$ is the parameter matrix of the i-th relation. Feature Encoding: The model encodes features along two spatial dimensions using different pooling kernels horizontal (1, H) and vertical (1, W). This results in two sets of features, each capturing information along one spatial dimension.

Horizontal encoding is formalized as:

$$z_r^h(h) = \frac{1}{|W|}\sum_{0<i<H} v_r^z(h,i) \tag{7}$$

Vertical encoding is formalized as:

$$z_r^w(w) = \frac{1}{|H|}\sum_{0<i<H} v_r^z(j,w) \tag{8}$$

Feature Aggregation and Transformation: The encoded features are concatenated and passed through a shared 1×1 convolutional transform function F_i, followed by a non-linear activation function δ, yielding intermediate features shown in equation (9).

$$f = \sigma(F_1([z^h, z^w])) \tag{9}$$

Feature Splitting and Further Transformation: The aggregated feature map f is split along the spatial dimension into two separate tensors, which are then transformed by additional 1×1 convolutional transform F_h and F_w to match the channel dimensions of the input:

$$\begin{aligned} g^h &= \sigma\left(F_h\left(f^h\right)\right) \\ g^w &= \sigma\left(F_w\left(f^w\right)\right) \end{aligned} \tag{10}$$

Here, σ is the sigmoid function used for generating attention weights.

Output Computation: The final output of the coordinate attention block is computed as a product of the input feature and the attention weights along both dimensions:

$$V_r = v_r(i,j) + g_r^h(i) + g_r^w(j) \tag{11}$$

Importance of Spatial Encoding: The coordinate attention mechanism allows for the capture of long-range dependencies and precise location information by separately aggregating features along two spatial dimensions. This approach differs significantly from the traditional global pooling used in channel attention, as it preserves detailed positional information, crucial for accurately locating objects of interest in vision tasks.

The coordinate attention block in the BiSDRTE framework represents an advanced method for capturing spatial structure and long-range dependencies in token tasks. By decomposing global pooling into two 1D feature encodings and applying careful transformations, the model effectively captures both holistic spatial features and maintains precise positional details, enhancing its ability to locate and process relevant information in complex overlapping patterns tasks, The P is defined with shown in equation (12).

$$p_r^i = \sigma\left(\begin{bmatrix} v_r^{s-k} + V_r^{s-k}(i,j) \\ \\ 1 \end{bmatrix}^\top W_r^i \begin{bmatrix} v_r^{o'-j} + V_r^{o'-j}(i,j) \\ \\ 1 \end{bmatrix}\right) \tag{12}$$

Here we select the biaffine model mainly due to its following two advantages. First, it maintains a matrix for each relation, which can model the characteristics of a relation accurately. Second, its probability computation mechanism allows it can accurately mine the interactions between a subject and an object. Both advantages are very helpful for improving the extraction precision.

$\mathcal{L}_r$ Loss to train the relation component, we also define a cross entropy-based loss, as shown in equation (13), where R is the predefined relation set and |R| is the number of total relations.

$$\mathcal{L}_r = \frac{1}{|R|}\sum_{i=1}^{|R|} \text{cew}\left(p_r^i, t_r^i\right) \tag{13}$$

F. SHARE-AWARE LEARNING MECHANISM

There are five extraction modules in BiSDRTE. During the multi-task learning-based training, each of them will form a relatively independent extraction task with the Encoder module. We use the popular teacher-forcing mode to train all the tasks except the ones that only take the original sentence as input. Under this mode, each task randomly selects some correct samples as input for training. Besides, to alleviate the exposure bias issue [24], we merge some randomly generated negative samples into the correct samples and use them together to train these tasks where the teacher forcing mode is used. The negative samples can simulate the real scenario in the inference phase, which helps train a robust model. Accordingly, the mentioned exposure bias issue is alleviated greatly. The loss function part combines the concept of residual network (RN) to optimize the loss function [32], and the specific calculation formula is detailed in formulas (6) and (13). Finally, the overall loss of BiSDRTE is defined with equation (14).

$$
\begin{aligned}
\mathcal{L}_1 &= \sum^{j\in \mathrm{L}} \mathcal{L}_j \\
\mathcal{L}_2 &= \sum^{j\in \mathrm{L}} \mathrm{w_i}\mathcal{L}_j \\
\mathrm{w_i} &= \frac{\mathcal{L}_\mathrm{i}}{|\mathcal{L}|} \\
\mathcal{L} &= \frac{\mathcal{L}_1 + \mathcal{L}_2}{2} \\
\mathrm{i} &= [1,2,3] \\
\mathrm{j} &= \left[\mathrm{S}_1, \mathrm{O}_2, \mathrm{O}'_1, \mathrm{S}_2, \mathrm{O}_1, \mathrm{S}_2, \mathrm{r}\right]
\end{aligned} \tag{14}
$$

However, we observe that the parameters in the shared Encoder module will receive back propagated gradients from the parameters of each extraction module. Consequently, the convergence rate of the Encoder module will be much different from those in other ex-traction modules. This will result in a convergence rate inconsistency issue, which means if we set a unified learning rate for these five extraction modules and the Encoder module, it would be difficult for them to converge to their optimal points simultaneously. In other words, some modules will be over-trained while others will be under-trained under a unified learning rate. So we propose a share-aware learning mechanism that assigns different learning rates for different modules. The basic idea of this mechanism is that the more tasks a module is shared, the smaller the learning rate it should be assigned. For example, the Encoder module should be assigned a smaller learning rate than other extraction modules since it is shared by more tasks. Specifically, the proposed learning mechanism assigns learning rates with equation (15).

$$
\xi_\mathrm{i} = \{^{\xi, \mathrm{k_i} \leq 1}_{\frac{(\delta)}{\mathrm{f(k_i)}} * \xi, \mathrm{k_i} > 1} \tag{15}
$$

where ξ is a base learning rate, ξ_i is the learning rate for the i-th module and $\mathrm{k_i}$ is the number of tasks that the i-th module is shared by. For example, in BiSDRTE, the corresponding k of the Encoder module would be 5 since this module is shared by all the five tasks, while the corresponding k of the subject tagger module in the sp2o direction would be 1 since this module is only used by its own task. $\delta \in [0, 1]$ is a regulatory factor that is used to finely adjust the learning rate, and f(.) to a reasonable real value (often larger than 1) so as to determine the major magnitude of the learning rate.

IV. EXPERIMENTS

A. DATASET AND DETAILDETAIL

In our study, we evaluated the BiSDRTE model on a variety of benchmark datasets: NYT [28], WebNLG [29], NYT10 [18], and NYT11 [12]. Following the precedent set by recent work [20], [23], [24], we utilized preprocessed versions of these datasets as released by [31] for NYT and WebNLG, and by [21] for NYT10 and NYT11 [3]. Table 2 provides a statistical overview of these datasets.

It's noteworthy that both NYT and WebNLG exist in two versions based on distinct annotation standards: 1) annotating only the last token of entities, and 2) annotating the entire span of entities. For clarity in our experiments, we refer to datasets following the first annotation standard as NYT∗ and WebNLG∗, and those following the second standard as NYT and WebNLG. The full-span annotated datasets offer a more comprehensive assessment of a model's true performance capabilities.

Additionally, it's observed [3] that NYT10 and NYT11 are less prominent compared to NYT or WebNLG. These are primarily employed to demonstrate a model's generalization ability, as most test sentences in them fall under the 'Normal' class. Therefore, for brevity, we only include them in our main experiment section.

In the discussed research, the evaluation of results employs standard micro precision, recall, and F1 score as key metrics. For the Relation Extraction Task (RTE), two distinct matching criteria are used:

1) PARTIAL MATCH

This criterion considers an extracted data triple (comprising a relation, subject entity, and object entity) correct if the predicted relation and the head (main component) of both the subject and object entities are correctly identified.

2) EXACT MATCH

Under this standard, a data triple is deemed correct only if the entire entities and the relation are completely and accurately matched with a known correct triple.

To maintain consistency with prior research (cited as [22], [25], [26]), Partial Match is applied to the NYT* and WebNLG* datasets, while Exact Match is used for the NYT and WebNLG datasets.

Regarding implementation details, the BiSDRTE model is trained using the AdamW optimizer. A threshold value of 0.5 is set for determining the presence of a subject, an object, or a relation. Specific values for certain parameters are defined, such as setting to, the regulatory factor to 0, and defining the mapping function as an identity function. Batch sizes are tailored to each dataset: 18 for NYT, NYT*, NYT10, and NYT11, and 6 for WebNLG and WebNLG*. These hyper-parameters are optimized based on developmental set outcomes, while other parameters are initialized randomly [7].

In the experiments, the BERT (base) model is utilized across all datasets. Each model is run five times, and the average of these runs is taken as the final reported result.

We have selected a range of robust and state-of-the-art models as baselines for our study. These include:

RIN (Sun et al., 2020) [21], $\text{PMEI}_{\text{LSTM}}$(Sun et al., 2021) [22], $\text{TPLinker}_{\text{LSTM}}$ (Wang et al., 2020) [25], $\text{R-BPtrNet}_{\text{LSTM}}$ (Chen et al., 2021) [7], $\text{CGT}_{\text{UniLM}}$(Ye et al., 2021) [28], $\text{CasRel}_{\text{BERT}}$ (Wei et al., 2020) [27], $\text{PMEI}_{\text{BERT}}$ (Sun et al., 2021) [22], $\text{TPLinker}_{\text{BERT}}$ (Wang et al., 2020) [25], $\text{StereoRel}_{\text{BERT}}$ (Tian et al., 2021) [24], $\text{PRGC}_{\text{BERT}}$ (Zheng et al., 2021) [36], $\text{R-BPtrNet}_{\text{BERT}}$ (Chen et al., 2021) [7], $\text{BiRTE}_{\text{LSTM}}$ (Ren et al., 2022) [4]. OneRel (Shang et al., 2022) [5], $\text{BiRTE}_{\text{BERT}}$ (Ren et al., 2022) [4], ERGM (Gao et al., 2023) [3], BRASK (Zhang et al., 2023) [2], RoleAttrTE (Yao et al., 2023) [1]. The majority of the baseline results are directly replicated from their respective original publications.

B. MAIN RESULTS AND EXPANDED ANALYSIS

In the expansive landscape of relational triple extraction, our model BiSDRTE emerges as a formidable contender, showcasing superior performance across diverse datasets. First, we evaluate the contributions of the proposed bidirectional extraction framework from two aspects. The first part focuses on comparisons and performance on public datasets compare other Single flow model and double flow model. The second part further verifies the core components of the model, analyzing the proportions of correct and incorrect extractions across different datasets.

First, we analyzed the main experimental results presented in Tables 1 and 2. Precision and Recall Dynamics: While BiSDRTE exhibits a slight dip in precision, its recall scores are notably higher. This trade-off is a strategic outcome of the bidirectional framework, which, while extracting some noise pairs and slightly impacting precision, significantly enhances the recall. This results in an overall improvement in the F1 score, a harmonious balance between precision and recall, which is a more holistic measure of a model's performance. Comparison with Established Models: An intriguing aspect of BiSDRTE's performance is its consistent outperformance of models like CasRel, TPLinker, and PRGC, showcasing superior entity and relation extraction capabilities, when utilizing the BERT encoder. The advanced capabilities of BERT, integrated into BiSDRTE's framework, contribute significantly to this enhanced performance.

Performance Across Datasets: BiSDRTE's excellence extends beyond standard datasets like NYT* and WebNLG*. Its prowess is equally evident in the NYT10 and NYT11 datasets. These datasets, often used to gauge a model's

TABLE 2. Statistics of datasets. N is the number of triples in a sentence.

Category	Dataset				Details of Test Set									
Category	Train	Valid	Test	Relations	Normal	SEO	EPO	HTO	N=1	N=2	N=3	N=4	N>5	Triples
NYT*	56,195	4,999	5,000	24	3,266	1,297	978	45	3,244	1,045	312	291	108	8,110
WebNLG*	5,019	500	703	171	245	457	26	84	266	171	131	90	45	1,591
NYT	56,195	5,000	5,000	24	3,222	1,273	969	117	3,240	1,047	314	290	109	8,120
WebNLG	5,019	500	703	216	239	448	6	85	256	175	138	93	41	1,607

generalization ability, further affirm BiSDRTE's robustness. The model's adaptability to different annotation standards and match criteria (Partial Match vs. Exact Match) underscores its potential for broad application. Special Mention of Competing Models: Notably, CGT (Cascaded Graph Transformer using UniLM), a model that leverages the unified language model (UniLM) for encoding sentences into a graph structure to enhance entity and relation extraction, and R-BPtrNet (Relation-Based Pointer Network employing extra entity type features), which incorporates additional entity type information to improve the precision of extraction, both demonstrate commendable performances. This highlights the innovative use of advanced NLP techniques and additional features in boosting the models' abilities to accurately identify and extract complex relationships within texts. However, BiSDRTE's unique bidirectional approach and learning mechanisms enable it to surpass these models in most scenarios. The ability of BiSDRTE to maintain high performance without relying on additional entity-type features, as in the case of R-BPtrNet, is particularly noteworthy. Enhanced Performance in Complex Scenarios: BiSDRTE's success is not just confined to standard scenarios but extends to complex sentence structures involving overlapping and multiple triples. The model's advanced architecture effectively addresses the challenges posed by such intricate data formats, a testament to its sophisticated design and engineering.

In summary, BiSDRTE's performance in our comprehensive evaluation, as depicted in Tables 3 and 4, is not just a testament to its technical prowess but also a promise of its applicability in real-world scenarios. The model adeptly balances precision and recall, excels in various datasets, and demonstrates a remarkable ability to handle complex sentence structures, setting a new standard in relational triple extraction.

Next, we provide a detailed description of the four main experimental setups to demonstrate the functionality of each component of the model and to examine the model's performance in correctly extracting or failing to extract triples.

1) EVALUATIONS OF COMPLEX SENTENCES

The performance of BiSDRTE in managing complex sentence structures, particularly those with overlapping triples (SEO) and multiple triples, is a crucial aspect of its robustness

TABLE 3. Main experiments, the main quote result uses of BiRTE (Ren et al., 2022).

Modal	Partial Match						Exact Match					
	NYT*			WebNLG*			NYT			WebNLG		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
RIN (2020)	87.2	87.3	87.3	87.6	87.0	87.3	83.9	85.5	84.7	77.3	76.8	77.0
PMEI$_{LSTM}$ (2021)	88.7	86.8	87.8	88.7	87.6	88.1	84.5	84.0	84.2	78.8	77.7	78.2
TPLinker$_{LSTM}$ (2020)	83.8	83.4	83.6	90.8	90.3	90.5	86.0	82.0	84.0	**91.9**	81.6	86.4
R-BPtrNet$_{LSTM}$(2021)	90.9	91.3	91.1	90.7	94.6	92.6	–	–	–	–	–	–
CGT$_{UniLM}$ (2021)	94.7	84.2	89.1	92.9	75.6	83.4	–	–	–	–	–	–
CasRel$_{BERT}$ (2020)	89.7	89.5	89.6	93.4	90.1	91.8	89.8	88.2	89.0	88.3	84.6	86.4
PMEI$_{BERT}$ (2021)	90.5	89.8	90.1	91.0	92.9	92.0	88.4	88.9	88.7	80.8	82.8	81.8
TPLinker$_{BERT}$ (2020)	91.3	92.5	91.9	91.8	92.0	91.9	91.4	92.6	92.0	88.9	84.5	86.7
StereoRel$_{BERT}$ (2021)	92.0	92.3	92.2	91.6	92.6	92.1	92.0	92.3	92.2	–	–	–
PRGC$_{BERT}$ (2021)	**93.3**	91.9	92.6	94.0	92.1	93.0	**93.5**	91.9	92.7	89.9	87.2	88.5
R-BPtrNe$_{tBERT}$ (2021)	92.7	92.5	92.6	93.7	92.8	93.3	–	–	–	–	–	–
BiRTE$_{LSTM}$(2022)	86.5	89.0	87.7	90.5	91.6	91.0	86.4	87.1	86.7	85.2	87.3	86.2
OneRel(2022)	91.3	90.5	90.9	93.8	91.4	92.6	91.1	90.4	90.8	90.5	88.2	89.4
BiRTE$_{BERT}$(2022)	92.2	93.8	93.0	93.2	94.0	93.6	91.9	93.7	92.8	89.0	89.5	89.3
ERGM(2023)	93.3	91.5	92.4	94.2	91.2	92.7	–	–	–	–	–	–
BRASK(2023)	93.0	91.5	92.2	94.8	92.2	93.5	–	–	–	–	–	–
RoleAttrTE(2023)	91.8	91.1	91.4	94.8	92.4	93.6	–	–	–	–	–	–
BiSDRTE	92.8	**94.2**	**93.5**	94.1	**94.6**	**94.4**	92.2	**94.0**	**93.1**	91.5	**90.0**	**90.7**

TABLE 4. Main experiments, the main quote result uses of BiRTE (Ren et al., 2022).

Modal	Partial Match						Exact Match					
	NYT10			NYT11			NYT10			NYT11		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
PMEI$_{LSTM}$(2021)	79.1	67.2	72.6	56.0	58.6	57.2	75.4	65.8	70.2	55.3	57.8	56.5
CasRel$_{BERT}$(2020)	77.7	68.8	73.0	50.1	58.4	53.9	76.8	68.0	72.1	49.1	56.4	52.5
StereoRel$_{BERT}$(2021)	80.0	67.4	73.2	53.8	55.4	54.6	–	–	–	–	–	–
PMEI$_{BERT}$(2021)	79.1	70.4	74.5	55.8	59.7	57.7	77.3	69.7	73.3	54.9	58.9	56.8
TPLinker$_{BERT}$ (2020)	78.9	71.1	74.8	55.9	60.2	58.0	78.5	68.8	73.4	54.8	59.3	57.0
BiRTE$_{LSTM}$(2022)	79.0	68.8	73.5	55.1	60.4	57.6	76.1	67.4	71.5	54.1	60.5	57.1
BiRTE$_{BERT}$(2022)	80.6	71.8	76.0	56.4	62.0	59.1	80.1	71.4	75.5	55.0	61.2	57.9
BiSDRTE	**80.9**	**72.3**	**76.4**	**55.9**	**63.8**	**59.6**	**80.3**	**71.9**	**75.8**	**54.6**	**62.1**	**58.1**

and efficacy. Table 5 provides a detailed analysis of BiSDRTE's performance compared to other notable models across various sentence categories, including Normal, EntityPairOverlap (EPO), SingleEntityOverlap (SEO), and HeadTailOverlap (HTO), as well as sentences containing varying numbers of triples ($T = 1$, $T = 2$, $T=3$, $T=4$, $T\geq5$). BiSDRTE's Performance in Complex Sentences: BiSDRTE exhibits a consistent performance advantage across all categories, but it especially excels in the SEO category. This can be attributed to its bidirectional framework's ability to handle complex triple associations effectively, a common challenge in sentences with overlapping entity pairs.

In the SEO category, BiSDRTE achieves a remarkable F1 score of 95.2% on NYT* and 96.2% on WebNLG*, indicating its superior capability to process sentences where a single entity is part of multiple relational triples.

- Comparison with Other Models: When compared to models like CasRel$_{BERT}$ and TPLinker$_{BERT}$, BiSDRTE not only shows better overall performance but also demonstrates a significant improvement in handling sentences with higher complexity ($T\geq5$). This underlines BiSDRTE's proficiency in extracting and managing a large number of relational triples within a single sentence. PRGC$_{BERT}$ and SPN models show competitive performance in certain categories, but BiSDRTE maintains an edge, particularly in the SEO and EPO categories. This is indicative of BiSDRTE's robust handling of sentences where traditional models might struggle due to overlapping entities and relations.
- Advantages in Handling Overlapping Entities: The SEO category, where entities overlap with multiple relations, is often challenging for RTE systems. BiSDRTE's advanced algorithms effectively disentangle these complex relationships, reflected in its high scores in this category. This is a significant advancement over traditional models, which may falter in accurately extracting multiple, overlapping triples. In the HTO and EPO categories, BiSDRTE's performance remains consistently high, showcasing its ability to correctly identify and extract triples even when entities are densely packed within a sentence.
- Performance Across Different Triple Counts: BiSDRTE's performance remains stable and superior across sentences with varying numbers of triples. This consistency is particularly notable in sentences with a higher count of triples ($T\geq5$), where the model's sophisticated

TABLE 5. F1 scores on sentences with different overlapping patterns and different triplet numbers (Normal, EntityPairOverlap (EPO), SingleEntityOverlap (SEO), and HeadTailOverlap (HTO)). Compare results of CasRel are copied from TPLinker directly. “T” is the number of triples contained in a sentence.

Modal	NYT*									WebNLG*								
	Normal	HTO	SEO	EPO	T = 1	T = 2	T = 3	T = 4	T≥5	Normal	HTO	SEO	EPO	T = 1	T = 2	T = 3	T = 4	T ≥ 5
$\mathrm{CasRel}_{\mathrm{BERT}}$	87.3	77.0	91.4	92.0	88.2	90.3	91.9	94.2	83.7	89.4	90.4	92.2	94.7	89.3	90.8	94.2	92.4	90.9
$\mathrm{TPLinker}_{\mathrm{BERT}}$	90.1	90.1	93.4	94.0	90.0	92.8	93.1	96.1	90.0	87.9	80.4	92.5	95.3	88.0	90.1	94.6	93.3	91.6
$\mathrm{PRGC}_{\mathrm{BERT}}$	91.0	81.8	94.0	94.5	91.1	93.0	93.5	95.5	93.0	90.4	94.6	93.6	95.9	89.9	91.6	95.0	94.8	92.8
SPN	90.8	-	94.0	94.1	90.9	93.4	**94.2**	95.5	90.6	-	-	-	-	89.5	91.3	**96.4**	94.7	93.8
$\mathrm{RBPtrNet}_{\mathrm{BERT}}$	90.4	-	94.4	**95.2**	89.5	93.1	93.5	**96.7**	91.3	89.5	-	93.9	**96.1**	88.5	91.4	96.2	**94.9**	94.2
$\mathrm{BiRTE}_{\mathrm{BERT}}$	91.4	-	94.7	94.2	91.5	93.7	93.9	95.8	92.1	90.1	-	95.9	94.3	90.2	92.9	95.7	94.6	92.0
Onerel	90.6	90.8	94.8	95.1	90.5	93.4	93.9	96.5	**94.2**	91.9	94.9	94.7	95.4	91.4	93.0	95.9	**95.7**	94.5
RoleAttrTE	89.6	-	93.6	93.1	89.9	91 .5	93.0	95.9	89.9	**93.9**	-	94.6	94.7	90.8	91.8	95.2	94.9	92.6
BiSDRTE	**91.2**	**91.3**	**95.2**	94.7	**92.2**	**93.9**	94.1	96.4	92.7	91.7	**95.1**	**96.2**	94.6	**91.9**	**93.4**	96.1	94.3	**94. 5**

TABLE 6. Results of detailed evaluations about model Partial Match and Exact Match of performance in NTY*, WebNLG*, NYT, WebNLG.

Modal	Partial Match						Exact Match					
	NYT*			WebNLG*			NYT			WebNLG		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
BiSDRTE	**92.8**	**94.2**	**93.5**	**94.1**	**94.6**	**94.4**	**92.2**	**94.0**	**93.1**	**91.5**	**90.0**	**90.7**
$\mathrm{BiSDRTE}_{\mathrm{sp2o}}$	91.6	92.1	91.8	93.0	93.4	93.2	92.0	91.9	92.0	89.3	89.7	89.5
$\mathrm{BiSDRTE}_{\mathrm{op2s}}$	91.0	92.8	92.0	92.7	93.7	93.2	91.2	92.4	91.8	88.7	90.2	89.4
$\mathrm{BiSDRTE}_{\mathrm{CA}}$	90.9	93.3	92.1	93.5	93.3	93.4	91.9	92.0	91.9	89.0	90.4	89.7
$\mathrm{BiSDRTE}_{\mathrm{RN}}$	92.0	92.8	92.4	94.0	93.5	93.7	92.9	91.9	92.9	89.4	90.9	90.1

architecture successfully navigates the increased complexity. This is a critical aspect for real-world applications where sentences can contain multiple, intricately linked entities and relations, necessitating a model that can handle such diversity without a drop in performance.

In summary, the analysis of complex sentence structures in Table 5, unequivocally positions BiSDRTE as a leading model in the realm of RTE. Its superior performance across various challenging categories, especially in sentences with overlapping and multiple triples, underscores its advanced design and practical applicability in diverse, real-world scenarios. The bidirectional framework, a cornerstone of BiSDRTE's design, proves instrumental in achieving these results, marking a significant step forward in the field of natural language processing and relational triple extraction.

2) DETAILED ABLATION EVALUATIONS

In the detailed evaluation of BiSDRTE, the Table 6 presents the ablation experiments on the main components of the model. In the detailed evaluation of BiSDRTE, as depicted in Table 6, we delved into several critical aspects of the model's performance, scrutinizing its efficacy across different frameworks and configurations. In Table 6, BiSDRTEsp2o, BiSDRTEop2s, BiSDRTECA, and BiSDRTERN are shown. Here, sp2o, op2s, CA, and RN represent the removal of the corresponding modules from the model.

- Bidirectional vs. Unidirectional Frameworks: The stark contrast in performance between BiSDRTE and the advantages of using s2o, o2s labeling methods to identify complex text structures and BIO (Beginning, Inside, Outside) labeling methods have been demonstrated in the literature [3]. Analyzing the ground entity extraction results, it becomes clear that the bidirectional framework of BiSDRTE significantly enhances the accuracy of entity extractions. The mutual reinforcement achieved by the intertwining of the s2o and o2s directions within the bidirectional framework results in a substantial improvement in both precision and recall, as reflected in the increased F1 scores. This synergy is absent in the unidirectional variants, leading to their diminished performance. The BiSDRTE 2-step variant, which employs a simplified 2-step extraction strategy, demonstrates a significant reduction in precision, indicating that while this approach may alleviate the ground entity extraction failure issue, it introduces a substantial amount of noise. This reinforces the conclusion that BiSDRTE's bidirectional, multifaceted approach is more effective in maintaining a balance between precision and recall.
- Coordinate Attention (CA), and Residual Network (RN): The dual-stream architecture, featuring sp2o and op2s streams, is a cornerstone of BiSDRTE's design, allowing it to capture relational information from both subject-to-object and object-to-subject perspectives. This comprehensive approach ensures that no aspect of the relationship context is overlooked. The CA mechanism integrated into BiSDRTE further refines its ability to discern intricate relationships between entity pairs. CA enables the model to focus more precisely on the relational aspects, leading to enhanced extraction accuracy. Moreover, the inclusion of RN plays a crucial role in preserving learned features while still incorporating original inputs. This blend of old and new insights

TABLE 7. F1 results of the ground entity extraction.

Models	Direction	NYT□	WebNLG□	NYT	WebNLG
BiSDRTE	sp2o	94.8	95.9	94.5	91.9
	op2s	94.4	96.0	93.5	91.2
BiSDRTEsp2o	sp2o	93.8	93.0	93.3	89.4
BiDRTEop2s	op2s	93.5	93.2	93.0	89.2

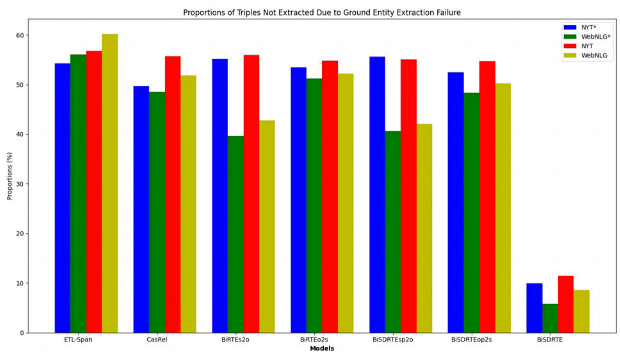

FIGURE 2. Proportions (%) of triples that are not extracted due to the ground entity extraction failure issue.

empowers BiSDRTE to adapt and learn more effectively, significantly boosting its performance.

In summary, the detailed analysis in Table 6 firmly establishes the superiority of BiSDRTE's bidirectional framework over unidirectional approaches and simpler pipeline variants. The bidirectional framework not only enhances entity extraction accuracy but also effectively addresses common extraction failure issues. This comprehensive evaluation, backed by empirical data, positions BiSDRTE as a robust and advanced solution in the realm of RTE, capable of tackling complex relational structures with high precision and recall.

3) THE GROUND ENTITY EXTRACTION

We evaluate the effectiveness of the proposed bidirectional extraction framework in extracting ground entities compared to unidirectional frameworks. To achieve this, we compare the ground entity extraction results of BiSDRTE, BiSDRTEsp2o, and BiDRTEop2s. The results are presented in Table 7. In the context of ground entity extraction results as shown in Table 7, the dual-stream architecture demonstrates its superiority. BiSDRTE's sp2o and op2s variants exhibit impressive F1 scores, outperforming other models in accurate entity pair extraction. These scores are a clear indication of the effectiveness of the dual-stream approach in capturing a more holistic view of entity relationships.

4) ANALYZING THE PROPORTIONS OF UNEXTRACTED TRIPLES DUE TO GROUND ENTITY EXTRACTION FAILURE.

We compared the proportions of unextracted triplets due to ground entity extraction failures between BiSDRTE and other tagging-based methods. This comparison demonstrates that our model performs better with more data.

BiSDRTE's advanced architecture markedly reduces such instances. Compared to models like BiRTE and CasRel, BiSDRTE demonstrates a significant reduction in extraction failures, further evidencing the robustness of its approach. Adaptability Evaluations BiSDRTE's adaptability is particularly evident in its handling of ground entity extraction failures. The model's bidirectional framework dramatically lowers the proportion of triples not extracted due to these failures, as illustrated in Figure 2. This reduction is a testament to the model's resilience and its ability to effectively manage complex extraction scenarios. The bidirectional approach ensures that even if one directional stream encounters difficulties, the other can compensate, thereby significantly reducing the likelihood of extraction failures.

In summary, the integration of a share-aware learning mechanism, dual-stream architecture, coordinate attention, and residual networks in BiSDRTE's culminates in a model that is not only proficient in relation extraction but also exhibits a high degree of adaptability and resilience. These features collectively establish BiSDRTE's as a state-of-the-art model in the realm of RTE, capable of navigating the complexities of relational contexts with unparalleled precision and efficiency.

V. CONCLUSION

This research introduces a pioneering RTE model addressing two key challenges in relation extraction: ground entity extraction failure and struggles with complex relational structures. Our bidirectional extraction framework effectively resolves the first issue, enhancing extraction accuracy. The share-dependent bidirectional Framework learning mechanism, complemented by coordinate attention and residual network integration, addresses the second, ensuring consistent learning across shared structures. Experiments across benchmarks confirm our model's state-of-the-art performance and robustness, highlighting its potential for advanced relational context understanding. This work marks a significant stride in RTE, showcasing the efficacy of integrating innovative architectural concepts for more precise extraction.

REFERENCES

[1] M. Lee, G. Yu, and G. Y. Li, "Graph embedding-based wireless link scheduling with few training samples," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2282–2294, Apr. 2021.

[2] L. Zhang, "Exploring self-distillation based relational reasoning training for document-level relation extraction," in *Proc. AAAI Conf. Artif. Intell.*, Jun. 2023, vol. 37, no. 11, pp. 13967–13975.

[3] Y. Yao, S. Mao, N. Zhang, X. Chen, S. Deng, X. Chen, and H. Chen, "Schema-aware reference as prompt improves data-efficient knowledge graph construction," in *Proc. 46th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2023, pp. 911–921.

[4] F. Ren, L. Zhang, X. Zhao, S. Yin, S. Liu, and B. Li, "A simple but effective bidirectional framework for relational triple extraction," in *Proc. 15th ACM Int. Conf. Web Search Data Mining*, Feb. 2022, pp. 824–832.

[5] Y. Shang, H. Huang, and X. Mao, "OneRel: Joint entity and relation extraction with one module in one step," in *Proc. 36th AAAI Conf. Artif. Intell. (AAAI)*, Jun. 2022, vol. 36, no. 10, pp. 11285–11293.

[6] G. Bekoulis, J. Deleu, T. Demeester, and C. Develder, "Joint entity recognition and relation extraction as a multi-head selection problem," *Expert Syst. Appl.*, vol. 114, pp. 34–45, Dec. 2018.

[7] Y. Chen, Y. Zhang, C. Hu, and Y. Huang, "Jointly extracting explicit and implicit relational triples with reasoning pattern enhanced binary pointer network," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics: Human Lang. Technol.*, 2021, pp. 5694–5703.

[8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*, vol. 1, 2019, pp. 4171–4186.
[9] L. Dong, N. Yang, and W. Wang, "Unified language model pre-training for natural language understanding and generation," in *Proc. 33rd Int. Conf. Neural Inf. Process. Syst.*, 2019, pp. 13063–13075.
[10] T. Dozat and C. D. Manning, "Deep biaffine attention for neural dependency parsing," in *Proc. Int. Conf. Learn. Represent.*, 2017, pp. 1–9.
[11] M. Eberts and A. Ulges, "Span-based joint entity and relation extraction with transformer pre-training," in *Proc. Frontiers Artif. Intell. Appl. ECAI*, vol. 325, 2020, pp. 2006–2013.
[12] T.-J. Fu, P.-H. Li, and W.-Y. Ma, "GraphRel: Modeling text as relational graphs for joint entity and relation extraction," in *Proc. 57th Annu. Meeting Assoc. Comput. Linguistics*, Florence, Italy, 2019, pp. 1409–1418.
[13] C. Gardent, A. Shimorina, S. Narayan, and L. Perez-Beltrachini, "Creating training corpora for NLG micro-planners," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, 2017, pp. 179–188.
[14] P. Gupta, H. Schtze, and B. Andrassy, "Table filling multi-task recurrent neural network for joint entity and relation extraction," in *Proc. COLING 26th Int. Conf. Comput. Linguistics, Tech. Papers*, 2016, pp. 2537–2547.
[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015, pp. 1–12.
[16] M. Miwa and M. Bansal, "End-to-end relation extraction using LSTMs on sequences and tree structures," in *Proc. 54th Annu. Meeting Assoc. Comput. Linguistics*, 2016, pp. 1105–1116.
[17] T. Nayak and H. T. Ng, "Effective modeling of encoder-decoder architecture for joint entity and relation extraction," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, Apr. 2020, pp. 8528–8535.
[18] T. Nayak and H. T. Ng, "Effective modeling of encoder–decoder architecture for joint entity and relation extraction," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 8528–8535.
[19] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1532–1543.
[20] S. Riedel, L. Yao, and A. McCallum, "Modeling relations and their mentions without labeled text," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*, Barcelona, Spain, Sep. 2010, pp. 148–163.
[21] K. Sun, R. Zhang, S. Mensah, Y. Mao, and X. Liu, "Recurrent interaction network for jointly extracting entities and classifying relations," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2020, pp. 3722–3732.
[22] K. Sun, R. Zhang, S. Mensah, Y. Mao, and X. Liu, "Progressive multitask learning with controlled information flow for joint entity and relation extraction," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 1–9.
[23] R. Takanobu, T. Zhang, J. Liu, and M. Huang, "A hierarchical framework for relation extraction with reinforcement learning," in *Proc. AAAI Conf. Artif. Intell.*, Jul. 2019, vol. 33, no. 1, pp. 7072–7079.
[24] X. Tian, L. Jing, L. He, and F. Liu, "StereoRel: Relational triple extraction from a stereoscopic perspective," in *Proc. 59th Annu. Meeting Assoc. Comput. Linguistics 11th Int. Joint Conf. Natural Lang. Process.*, 2021, pp. 4851–4861.
[25] Y. Wang, B. Yu, Y. Zhang, T. Liu, H. Zhu, and L. Sun, "TPLinker: Single-stage joint extraction of entities and relations through token pair linking," in *Proc. 28th Int. Conf. Comput. Linguistics*, 2020, pp. 1572–1582.
[26] S. Zheng, F. Wang, H. Bao, Y. Hao, P. Zhou, and B. Xu, "Joint extraction of entities and relations based on a novel tagging scheme," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, 2017, pp. 1227–1236.
[27] Z. Wei, J. Su, Y. Wang, Y. Tian, and Y. Chang, "A novel cascade binary tagging framework for relational triple extraction," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics*, 2020, pp. 1476–1488.
[28] H. Ye, N. Zhang, and S. Deng, "Contrastive triple extraction with generative transformer," in *Proc. AAAI Conf. Artif. Intell.*, 2021, vol. 35, no. 16, pp. 14257–14265.
[29] B. Yu, Z. Zhang, X. Shu, T. Liu, Y. Wang, B. Wang, and S. Li, "Joint extraction of entities and relations based on a novel decomposition strategy," in *Proc. ECAI*, 2019, pp. 2282–2289.
[30] Y. Yuan, X. Zhou, S. Pan, Q. Zhu, Z. Song, and L. Guo, "A relation-specific attention network for joint entity and relation extraction," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, vol. 4, Jul. 2020, pp. 4054–4060.
[31] D. Zeng, H. Zhang, and Q. Liu, "CopyMTL: Copy mechanism for joint extraction of entities and relations with multi-task learning," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 9507–9514.
[32] K. He, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
[33] X. Zeng, S. He, D. Zeng, K. Liu, S. Liu, and J. Zhao, "Learning the extraction order of multiple relational facts in a sentence with reinforcement learning," in *Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP)*, 2019, pp. 367–377.
[34] X. Zeng, D. Zeng, S. He, K. Liu, and J. Zhao, "Extracting relational facts by an end-to-end neural model with copy mechanism," in *Proc. 56th Annu. Meeting Assoc. Comput. Linguistics*, 2018, pp. 506–514.
[35] M. Zhang, Y. Zhang, and G. Fu, "End-to-end neural relation extraction with global optimization," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2017, pp. 1730–1740.
[36] H. Zheng, R. Wen, X. Chen, Y. Yang, Y. Zhang, Z. Zhang, N. Zhang, B. Qin, X. Ming, and Y. Zheng, "PRGC: Potential relation and global correspondence based joint relational triple extraction," in *Proc. 59th Annu. Meeting Assoc. Comput. Linguistics 11th Int. Joint Conf. Natural Lang. Process.*, 2021, pp. 6225–6235.
[37] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2021, pp. 13713–13722.

YOUREN CHEN received the B.S. degree in network engineering from Xinjiang Normal University, where he is currently pursuing the degree. His current research interests include deep learning, intelligent systems, and knowledge graphs.

YONG LI received the Ph.D. degree in computer science from Nanjing University of Aeronautics and Astronautics, in 2018. He is currently an Associate Professor with Xinjiang Normal University. His research interests include machine learning and intelligent software engineering.

MING WEN received the degree in automatic control from Xi'an University of Technology, in 1988. He is currently the Director and a Researcher with the Software Development and Testing Center, Xinjiang Electronic Research Institute. His research interests include software engineering and artificial intelligence.