

A Domain Question Answering Algorithm Based on the Contrastive Language-Image Pretraining Mechanism

Zuoxing Zhang, Dong Liang, Zhen Zhang, Yang Cai, Hongyi Hou

Nanjing Research Institute of Electronic Engineering, Nanjing, China

Email: zhangzuoxing1993@163.com

How to cite this paper: Zhang, Z.X., Liang, D., Zhang, Z., Cai, Y. and Hou, H.Y. (2023) A Domain Question Answering Algorithm Based On the Contrastive Language-Image Pretraining Mechanism. *Journal of Computer and Communications*, 11, 1-15.

<https://doi.org/10.4236/jcc.2023.115001>

Received: April 10, 2023

Accepted: May 8, 2023

Published: May 11, 2023

Copyright © 2023 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Research on specific domain question-answering technology has become important with the increasing demand for intelligent question-answering systems. This paper proposes a domain question-answering algorithm based on the CLIP mechanism to improve the accuracy and efficiency of interaction. First, this paper reviewed relevant technologies involved in the question-answering field. Then, the question-answering model based on the CLIP mechanism was produced, including its design, implementation, and optimization. It also described the construction process of the specific domain knowledge graph, including graph design, data collection and processing, and graph construction methods. The paper compared the performance of the proposed algorithm with classic question-answering algorithms BiDAF, R-Net, and XLNet models, using a military domain dataset. The experimental results show that the proposed algorithm has advanced performance, with an F1 score of 84.6% on the constructed military knowledge graph test set, which is at least 1.5% higher than other models. We conduct a detailed analysis of the experimental results, which illustrates the algorithm's advantages in accuracy and efficiency, as well as its potential for further improvement. These findings demonstrate the practical application potential of the proposed algorithm in the military domain.

Keywords

QA system, Knowledge Graph, Domain Knowledge, Knowledge Retrieval

1. Introduction

With the rapid development of information technology, the need for acquiring and processing massive amounts of information has become increasingly urgent.

Among them, question-answering systems [1] have become an important human-computer interaction method and are receiving more and more attention from academia and industry. As an important branch of natural language processing, domain question-answering technology is an effective way to help people quickly obtain the desired information and has received widespread attention and research. However, existing domain question-answering technology [2] still faces many issues, such as inadequate understanding and utilization of domain knowledge, model accuracy, and efficiency, which limits its further promotion and application in practical use. Therefore, this study conducted in-depth research on domain question-answering technology based on the CLIP (Contrastive Language-Image Pre-Training, CLIP) mechanism [3], aiming to improve the accuracy and efficiency of domain question-answering and provide better solutions for its application.

The main issue addressed in this study is how to improve the accuracy and efficiency of domain question-answering technology and achieve accurate understanding and utilization of domain knowledge. The significance of this study lies in proposing a new domain question-answering technology that uses the CLIP mechanism to improve the accuracy and efficiency of domain question-answering and provide better solutions for practical applications. The main approach is to use the CLIP mechanism as the basic framework of domain question-answering technology, build a new type of domain question-answering model, and through experimental analysis, improve the model's accuracy and efficiency, and use domain knowledge graphs [4] to provide more accurate domain knowledge for the model, further improving the model's effectiveness and performance. This paper proposes a new domain question-answering framework, achieves accurate understanding and utilization of domain knowledge, and provides better solutions for practical applications.

The structure of this paper is as follows: first, it introduces the relevant technologies of domain question-answering, proposes research on domain question-answering technology based on the CLIP mechanism, including model design, implementation and optimization, and experimental result analysis. Secondly, it introduces the construction of domain knowledge graphs, including graph design, data collection and processing, graph construction algorithms, and experimental result analysis. Then, the experimental results are analyzed in detail, including question-answering effect analysis, knowledge graph construction effect analysis, and model optimization effect analysis. Finally, the research results and significance of this paper are summarized, and future work is outlined. The research approach and experimental results of this paper have certain reference value and are expected to provide some inspiration and guidance for the research and application of domain question-answering technology.

2. Related Work

In practical applications, domain question-answering (QA) technology is widely used in fields such as intelligent customer service [5], intelligent search [6], etc.

This chapter will introduce the relevant technologies of domain QA, including natural language processing, QA systems, knowledge graphs, and other aspects.

1) Natural language processing (NLP)

NLP refers to the technology that computers use to analyze and process natural language to achieve functions such as text understanding [7], generation [8], translation, etc. NLP includes aspects such as word segmentation, part-of-speech tagging, semantic analysis, and syntax analysis. In domain QA technology, NLP technology is widely used to analyze and understand user questions, in order to perform subsequent processing and provide answers.

2) QA systems

QA systems are systems based on natural language processing and knowledge representation, used to answer questions asked by users. QA systems are mainly divided into two types: pattern-matching QA systems [9] and semantic analysis QA systems [10]. The former mainly uses rule matching to find answers to questions, while the latter analyzes and understands the questions semantically, matches and infers them with relevant knowledge in the knowledge base, and obtains the answers.

3) Knowledge graphs

Knowledge graphs represent entities, properties, and relationships as nodes and edges, forming a large graphical structure. They mainly include knowledge representation, knowledge storage, and knowledge inference, and are widely used in domain QA to accurately understand and utilize domain knowledge.

4) Domain QA

Currently, deep learning-based QA technology has become a hot research topic in the domain QA field. Typical models include convolutional neural networks, recurrent neural networks, attention mechanisms, etc. These models improve the accuracy and generalization performance of the model through learning and training on large amounts of data. In addition, some studies use knowledge graphs and graph neural networks to achieve deep learning representation and inference of domain knowledge, improving the accuracy and intelligence of QA. Furthermore, to improve the performance of domain QA technology, techniques such as data augmentation [11] [12], multi-task learning [13], and semi-supervised learning can enhance the robustness and generalization ability of the model from different perspectives, and reduce data annotation costs. The development of technologies such as natural language processing, QA systems, knowledge graphs, and deep learning provides a solid foundation for the application of domain QA. This paper proposes a domain QA technology based on the CLIP mechanism on the basis of these technologies, aiming to improve the accuracy and intelligence of QA, and achieve a deep understanding and utilization of domain knowledge.

3. Construction of Domain Graph

In domain question answering, the construction of a knowledge graph plays an

important role in improving the effectiveness of question answering. The domain question answering based on the CLIP mechanism proposed in this paper requires a complete domain knowledge graph. **Figure 1** shows the basic process of constructing a domain knowledge graph. First, the domain knowledge base needs to be obtained and processed to remove redundant and erroneous information. Then, it needs to be transformed into a graph form, and some algorithms and strategies are used to optimize the structure and performance of the graph. Finally, the graph is combined with the CLIP-based question answering technology to achieve more accurate question answering services and more intelligent information interaction. During the construction and optimization of the graph, methods such as node clustering [14], relationship extraction, and

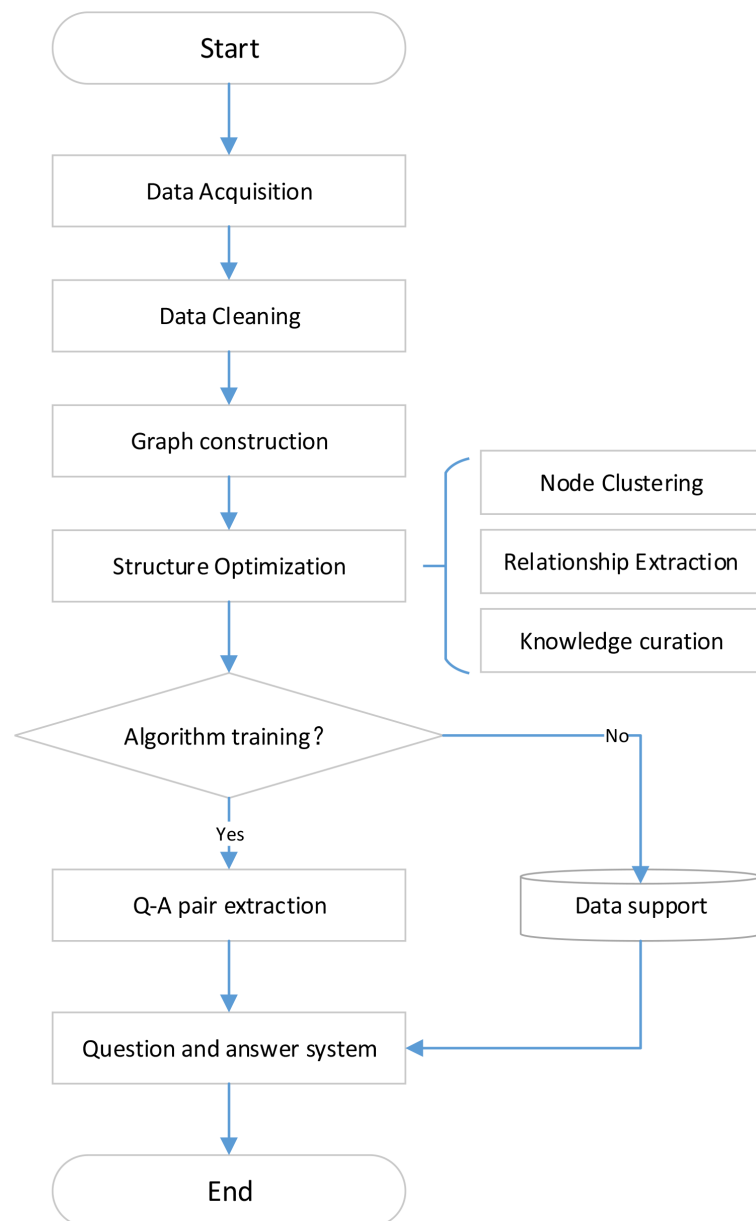


Figure 1. Domain knowledge base construction and question answering process.

Table 1. Data sources for domain knowledge base.

Forum	Content	Participants	Country
DefenceTalk	Defense technology and international military affairs	Military experts, professionals	USA
Military.com	Military news, opinions, job listings	Professionals, enthusiasts	USA
Warzone	Military and defense technology	Professionals, researchers	USA
RusMilitary	Russian and world military affairs	Professionals, researchers	Russia
m.de	German military and defense technology	Professionals, researchers, enthusiasts	Germany
DefenceTalk	European and international military affairs	Experts, professionals	Europe
JointForces	European and international military affairs	Professionals, researchers, government officials	Europe

knowledge inference [15] were used to optimize the structure and performance of the graph. In CLIP-based question answering, entities and relationships in the graph are used as a knowledge base, and user questions are transformed into query statements to find matching answers in the graph.

3.1. Construction of Domain Knowledge Base

This article aims to construct a military domain knowledge graph based on open forum data. First, we determined the entity and relationship types, including entities such as people, weapons and equipment, military organizations, and war events, and relationships such as command, ownership, and combat. After determining the entity and relationship types, we extracted a large amount of text, image, and video data from open military forums, as shown in **Table 1**, and represented the relationships between entities through knowledge graph triplets (head entity, relationship, tail entity).

In the data collection and processing section, the following types of data were extracted from the data sources listed in **Table 1**:

- 1) Military event data: event data related to military actions, wars, and military training, including event name, date, location, participants, and results.
- 2) Military equipment data: detailed parameters of various weapons and military equipment, such as range, power, and accuracy, as well as manufacturer and model information.
- 3) Military organization data: the organizational structure of military forces in various countries and regions, information on military commanders and leaders, and the relationships between various organizations.
- 4) Military history data: historical data on wars, conflicts, and military actions, as well as information on historical events such as people, places, and dates.
- 5) Military theory data: theoretical and research data on strategy, tactics, and weapon use, as well as data on military technology and development.

3.2. Graph Structure Optimization

After obtaining a reliable domain knowledge base, it was transformed into a graph structure, representing entities and relationships as nodes and edges.

Then, some algorithms and strategies were used to optimize the structure and performance of the graph, making it more compact and effective.

The methods used are as follows:

- 1) Node clustering: similar nodes were grouped together to reduce the size and complexity of the graph.
- 2) Relationship extraction: relationship information was extracted from the text using natural language processing techniques and added to the graph.
- 3) Knowledge inference: new information was inferred from the known information in the graph to enrich its content and knowledge.

Through the above process, a knowledge graph with millions of entities and relationships was eventually formed. The entities in the graph include people, organizations, events, equipment, history, and theory, and the relationships include commanders, participants, time, and location. In the process of building and optimizing the graph, entity linking and relationship extraction algorithms were used to ensure the accuracy and completeness of the data.

4. Intelligent Question-Answering Algorithm Design

The main feature of the CLIP model is the use of joint training to learn representations of both annotated images and text. This joint training approach enables the model to perform well on multiple tasks, including classification, semantic similarity, and question-answering. This paper used the CLIP model and graph neural networks (GNN) [16] as the core components of a question-answering system, and fine-tuned them to achieve domain-specific question-answering capabilities. A domain-specific question-answering algorithm was constructed based on the CLIP model and GNN, which utilized joint training to learn representations of annotated images and text. Supervised learning methods were used to fine-tune and transfer the model based on a military domain question-answering dataset. In this process, the BERT model [17] was used to process entity features, the ViT model [18] to process associated image features, and cross-entropy loss was used as the training objective. Regularization techniques were also used to prevent overfitting. Finally, the optimal answer was selected as the output by computing the similarity between the question vector and the answer vector. The structure of the domain-specific question-answering algorithm based on the CLIP mechanism was shown in **Figure 2**.

First, the constructed problem-answer pairs consist of general natural language questions and text answers associated with image information. The natural language questions and text answers are processed through a graph neural network (GNN) model for knowledge graph small-sample transfer learning. The GNN model is utilized for question subject processing and the question-answer pairs are built based on the knowledge graph. The GNN model is used for representation learning and the small-sample transfer learning method is employed to train the model with data from the knowledge graph, enabling the model to use existing information from the knowledge graph to enhance task performance.

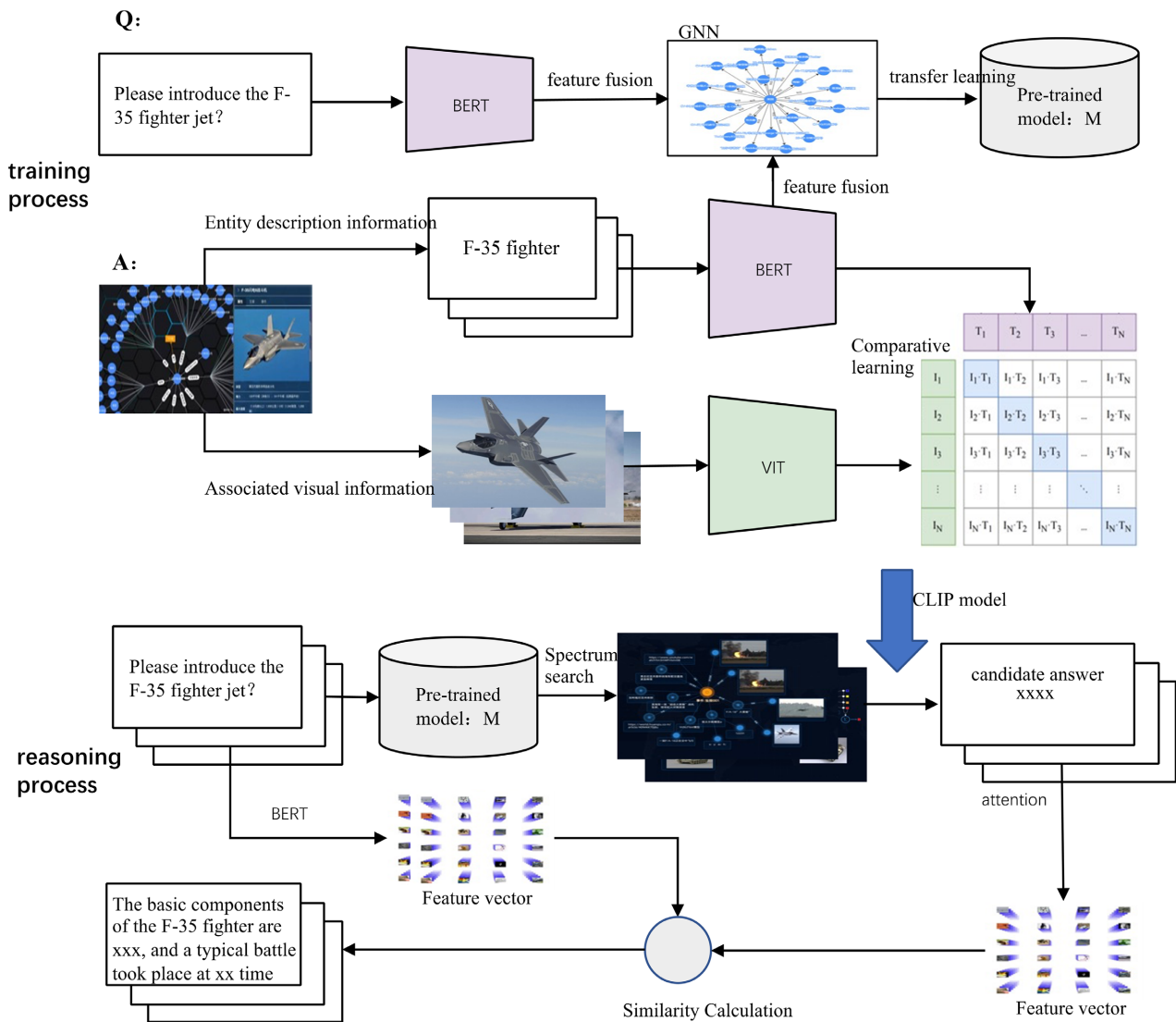


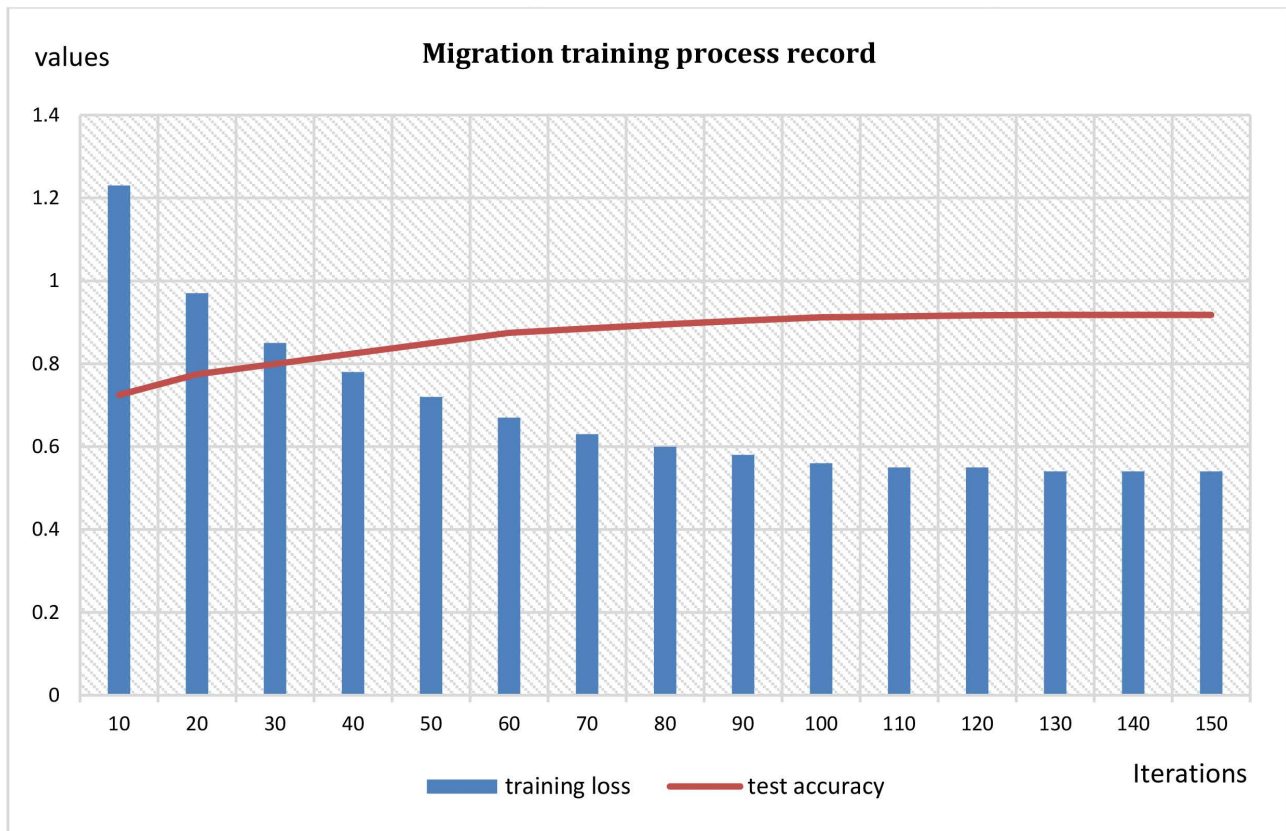
Figure 2. Question answering algorithm architecture based on CLIP mechanism.

Second, during the fine-tuning process of the CLIP model, a supervised learning method is used with the annotated problem-answer pairs in the knowledge graph for training. The questions and answers are treated as sequence data, and the BERT model is used to process entity features while the VIT model is used to process associated image features. The cross-entropy loss function is used as the training objective, minimizing the prediction error of the model on the training set. Regularization techniques such as dropout and weight decay are also used to prevent overfitting.

Then, a pre-extracted military domain question-answer dataset consisting of 500 question-answer pairs and 1000 images with semantic information is used as the training data. Pre-training Data and Methods are shown in **Table 2**, and the training records are shown in **Figure 3**. The Adam optimizer is used during the training process with a learning rate of 0.001, and the training stops when the model's accuracy on the validation set no longer improves. In each epoch, 5

Table 2. Pre-training data and methods.

Model	Data Source	Data Type	Data Volume	Training Method
GNN + CLIP	Domain Knowledge Graph	Text, Image	1500	Transfer Learning

**Figure 3.** Records the trend of the training process.

random questions and their corresponding answers and 5 images are selected as training samples from the question-answer data. These samples are fed into the CLIP model for feature extraction. To enable the model to utilize domain knowledge graphs for transfer learning, the entity and relationship information from the knowledge graph is fused with the training data as input during the training process.

Finally, during the inference process, the transferred CLIP model is used to extract and retrieve information from the knowledge graph, transforming image information associated with entities into semantic features. This extends the richness of the candidate answers and facilitates the organization and arrangement of candidate answers. The final answer is obtained by calculating the similarity between the feature vectors of the question and the candidate set, completing the entire question-answering process.

In addition, in the semantic feature processing of the above question-answering system, we used two models: one for question representation and the other for answer representation. Specifically, the question representation model takes a

natural language question as input and outputs a vector representation that contains the semantic and syntactic information of the question. The answer representation model takes a candidate answer as input and outputs a vector representation that contains the semantic and syntactic information of the answer. By calculating the similarity between the question vector and the answer vector, the optimal answer can be selected as output. For the question representation model, we used the BERT model to take the question as a sequence input and obtain the vector representation of the question through multiple layers of neural networks. Specifically, assuming that the question is $Q = w_1, w_2, \dots, w_n$, where n is the length of the question, and w_i is the i -th word. We convert each word in the question into its corresponding word vector, and then add the word vectors to obtain the representation vector of the entire question:

$$q = \sum_{i=1}^n w_i \quad (1)$$

For the answer representation model, we use an attention mechanism to highlight the parts of the answer related to the question. Specifically, assuming that the answer is $A = w_1, w_2, \dots, w_m$, where m is the length of the answer, and w_i is the i -th word. We first convert each word in the answer into its corresponding word vector, and then calculate the similarity between each word vector and the question representation vector to obtain an attention distribution vector a . Finally, we weight the word vectors in the answer and sum them to obtain the representation vector of the entire answer:

$$a_i = \frac{\exp(q \cdot w_i)}{\sum_{j=1}^m \exp(q \cdot w_j)} \quad (2)$$

$$a = \sum_{i=1}^m a_i w_i \quad (3)$$

5. Experimental Design for QA

This experiment will use the military domain dataset constructed in the aforementioned paper, which includes 5000 questions and their corresponding answers, manually annotated and reviewed. Additionally, we will also conduct experiments on the SQuAD [19] dataset, which contains a broader range of domain knowledge and more types of questions, enabling better evaluation of the model's generalization ability.

5.1. Model Structure

This section will introduce the domain QA technology based on the CLIP mechanism proposed in this paper and compare it with three other classic QA models (BiDAF [20], R-Net [21], and XLNet [22]) to verify the effectiveness and superiority of the proposed algorithm. The experiment will be conducted on the military domain dataset constructed in the aforementioned paper and on the SQuAD dataset. BiDAF, R-Net, and XLNet models will be used to train the QA

models. The architectures and parameter settings of the three models are as follows in **Table 3**.

For each model, we will train on the training and validation sets, and use early stopping to prevent overfitting. The optimizer used for training is Adam, and the loss function is cross-entropy loss.

5.2. Dataset Construction

This experiment will use two datasets: the military domain dataset and the SQuAD dataset. The military domain dataset was extracted from the open-source military forum mentioned earlier and contains a large number of military QA data, with a total of 5000 question-answer pairs. The SQuAD dataset contains a large number of question-answer pairs from Wikipedia articles, with a total of 10,000 question-answer pairs.

5.3. Data Preprocessing

In the data preprocessing stage, the following steps need to be taken for both datasets: data cleaning, word segmentation, word vectorization, question-answer matching, etc.; randomly divide the dataset into training set, validation set, and test set; for each question, generate a candidate answer set, including the correct answer and some distractor answers. These steps aim to transform the raw text into a format that can be input to the model and improve the quality and accuracy of the input to the model.

5.4. Model Tuning

This experiment will compare four models: the domain question answering model based on the CLIP mechanism proposed in this paper, the BiDAF model,

Table 3. Comparison of model selection and algorithm structure.

Algorithm Model	Composition
BiDAF Model	<ul style="list-style-type: none">• Two bidirectional LSTM layers• Self-attention mechanism• Global attention mechanism• Predict the start and end positions of the answer
R-Net Model	<ul style="list-style-type: none">• Three bidirectional LSTM layers• Bitwise attention mechanism• Gated recurrent unit• Predict the start and end positions of the answer
XLNet Model	<ul style="list-style-type: none">• Neural network architecture based on Transformer• Use autoregressive mechanism to predict answers• Permutation-based language model

the R-Net model, and the XLNet model. These models will be trained on the military domain dataset and the SQuAD dataset. In the model training and tuning stage, the model needs to be tuned for hyperparameters, model structure adjustments, model weight initialization, etc.

5.5. Model Evaluation

In the model evaluation stage, metrics such as accuracy, recall, and F1 score will be used to evaluate the models. These metrics can measure the accuracy and completeness of the models in the question-answer matching task. By comparing the performance of each model on the two datasets and recording and analyzing the data generated during the experiment, the performance of the proposed algorithm, BiDAF model, R-Net model, and XLNet model on the military domain dataset and SQuAD dataset were compared, and the experimental results are shown in **Table 4**.

It can be seen that the proposed algorithm achieved the highest test F1 score of 84.6% on the military domain dataset, outperforming the other three models. On the SQuAD dataset, the test F1 score of the proposed algorithm was 78.2%, slightly lower than that of XLNet but higher than R-Net and BiDAF. In terms of test accuracy and recall, the proposed algorithm performed well compared to the other three models, with test accuracy and recalls both exceeding 80% on the military domain dataset.

In addition, we analyzed the performance of each model on different question types. On the constructed military domain dataset, the model performed best on “when” and “where” type questions. This means that the model has good processing capabilities for time, location, and parameter determination issues in the military domain. On the SQuAD dataset, the model performed slightly worse than the XLNet model on “what” and “how” type questions. This indicates that the performance of this model may be slightly inferior to other models for general reading comprehension tasks. Overall, the model demonstrated excellent performance in a specific field and has some generalization ability, some typical

Table 4. Comparison of experimental results.

Model	Dataset	Training Loss	Test Accuracy	Test Recall	Test F1 Score
Proposed	Military domain data	0.003	86.20%	83.10%	84.60%
BiDAF l	Military domain data	0.002	80.70%	77.30%	78.90%
R-Net	Military domain data	0.001	81.50%	78.60%	79.90%
XLNet	Military domain data	0.001	84.30%	81.90%	83.10%
Proposed	SQuAD dataset	0.002	80.50%	76.40%	78.20%
BiDAF	SQuAD dataset	0.004	76.20%	72.80%	74.30%
R-Net	SQuAD dataset	0.003	77.80%	74.20%	75.90%
XLNet	SQuAD dataset	0.005	82.40%	78.90%	80.60%

questions and answers are shown in **Table 5**.

5.6. Experimental Analysis

According to the above experimental results, our proposed CLIP-based model performs better than the BiDAF and R-Net models in military domain question answering tasks but slightly worse than the XLNet model in the open domain. We analyze as follows.

Firstly, the XLNet model in Reference [22] adopts a more complex pre-training model, which can better capture the semantic information of context when processing natural language processing tasks. Our proposed CLIP-based model is trained based on a multimodal pre-training model, which performs well in multimodal feature extraction but may not be as good as specialized natural language pre-training models for natural language processing.

Secondly, the special properties of the military domain may have an impact on the model's performance. The vocabulary and knowledge points involved in the military domain are more specialized and complex, requiring a deeper understanding and analysis of domain knowledge, which may still pose a challenge for the CLIP-based model.

Finally, the choice of evaluation metrics may have an impact on the experimental results. In our experiments, we chose common accuracy as the evaluation

Table 5. Typical types of problems.

Question	Answer
What is the primary weapon used by Russia's Su-57 fighter jet?	30 mm cannon and various missiles, such as air-to-air, air-to-ground, and anti-ship missiles.
How many variants of the F-35 fighter jet are there in the US?	There are three: F-35A (Air Force version), F-35B (short takeoff and vertical landing version), and F-35C (carrier-based version).
What missiles can be used by the US F-35 fighter jet?	AIM-9X Sidewinder air-to-air missile, AIM-120 AMRAAM medium-range air-to-air missile, AGM-154 JSOW air-to-ground missile, and AGM-158 JASSM long-range air-to-ground missile, etc.
What is Israel's "Iron Dome" system used for?	It is used to intercept low-altitude targets such as short-range rockets, mortar shells, and unmanned aerial vehicles.
What is the US Marine Corps' "Wasp" class amphibious assault ship?	It is a vessel used for maritime assault, transport, and support.
What is the maximum range of Russia's "Iskander-M" missile?	It can hit targets up to 550 km away.
How many generations of the UK Royal Air Force's "Typhoon" fighter jet are there?	There are three generations: F.2, F.3, and F.6.
Why was the Soviet T-34 tank widely used in World War II?	It was lightweight, agile, heavily armored, and had powerful firepower.
What missions can the US "Raptor" fighter jet perform?	It can perform multiple missions, including air superiority, ground attack, electronic warfare, intelligence gathering, and joint command.
What type of shell does the US M1A1 main battle tank use?	DU armor-piercing shells.

metric. However, in practical applications, different tasks may require different evaluation metrics. Therefore, different models may perform differently for different evaluation metrics.

6. Conclusion and Prospects

In this paper, we proposed a military domain question answering technology based on the CLIP mechanism, which, after sufficient data collection and processing, can construct a powerful domain knowledge graph, improving the efficiency of question answering and knowledge graph construction. In the experimental part, we compared the performance of our algorithm with the Bi-DAF, R-Net, and XLNet models on the military domain dataset, and the results showed that our algorithm achieved excellent performance in question answering. In future research, we will explore the following directions: first, we will further improve the question answering model and knowledge graph construction algorithm. Existing algorithms still have deficiencies in answering complex questions, and the graph construction algorithm needs further optimization. Secondly, we will explore more extensive field applications. Our algorithm performs well in the military domain, and we will explore its application to other fields, such as healthcare and finance. Finally, we will explore methods for model optimization. In addition to improving the model's accuracy, we will also study how to improve the model's operating efficiency and stability to better meet the needs of practical applications. The CLIP-based domain question answering technology proposed in this paper has good application prospects in the military domain, and future research directions will focus on introducing cross-modal feature extraction methods to further improve the algorithm and expand the application mode.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Kumar, A.J., Schmidt, C. and Köhler, J. (2017) A Knowledge Graph Based Speech Interface for Question Answering Systems. *Speech Communication*, **92**, 1-12. <https://doi.org/10.1016/j.specom.2017.05.001>
- [2] Kapanipathi, P., Abdelaziz, I., Ravishankar, S., *et al.* (2020) Question Answering over Knowledge Bases by Leveraging Semantic Parsing and Neuro-Symbolic Reasoning. (Preprint)
- [3] Radford, A., Kim, J.W., Hallacy, C., *et al.* (2021) Learning Transferable Visual Models from Natural Language Supervision. *Proceedings of the 38th International Conference on Machine Learning*, Virtual Event, 18-24 July 2021, 8748-8763.
- [4] Muralidhar, N., Islam, M.R., Marwah, M., *et al.* (2018) Incorporating Prior Domain Knowledge into Deep Neural Networks. *Proceedings of 2018 IEEE International Conference on Big Data (Big Data)*, Seattle, 10-13 December 2018, 36-45. <https://doi.org/10.1109/BigData.2018.8621955>

- [5] Zhang, B., Lin, H., Zuo, S., *et al.* (2020) Research on Intelligent Robot Engine of Electric Power Online Customer Services Based on Knowledge Graph. *Proceedings of the 2020 the 4th International Conference on Innovation in Artificial Intelligence*, Xiamen, 8-11 May 2020, 216-221. <https://doi.org/10.1145/3390557.3394312>
- [6] Dunn, M., Sagun, L., Higgins, M., *et al.* (2017) Searchqa: A New Q&A Dataset Augmented with Context from a Search Engine. (Preprint)
- [7] Zhang, Z., Geng, X., Qin, T., *et al.* (2021) Knowledge-Aware Procedural Text Understanding with Multi-Stage Training. *Proceedings of the Web Conference 2021 (WWW'21)*, Ljubljana Slovenia, 19-23 April 2021, 3512-3523. <https://doi.org/10.1145/3442381.3450126>
- [8] Yu, W., Zhu, C., Li, Z., *et al.* (2022) A Survey of Knowledge-Enhanced Text Generation. *ACM Computing Surveys*, **54**, 1-38. <https://doi.org/10.1145/3512467>
- [9] Hou, X., Zhu, C., Li, Y., *et al.* (2022) Question Answering System Based on Military Knowledge Graph. *International Conference on Electronic Information Engineering and Computer Communication (EIECC 2021)*, Nanchang, 17-19 December 2021, 33-39. <https://doi.org/10.1117/12.2634559>
- [10] Gu, Y., Pahuja, V., Cheng, G., *et al.* (2022) Knowledge Base Question Answering: A Semantic Parsing Perspective. (Preprint)
- [11] Yang, S., Xiao, W., Zhang, M., *et al.* (2022) Image Data Augmentation for Deep Learning: A Survey. (Preprint)
- [12] Bayer, M., Kaufhold, M.A. and Reuter, C. (2022) A Survey on Data Augmentation for Text Classification. *ACM Computing Surveys*, **55**, 1-39. <https://doi.org/10.1145/3544558>
- [13] Xia, J., Li, X., Tan, Y., *et al.* (2022) Event Detection via Context Understanding Based on Multi-Task Learning. Association for Computing Machinery, New York. <https://doi.org/10.1145/3529388>
- [14] Mustafi, D., Mustafi, A. and Sahoo, G. (2022) A Novel Approach to Text Clustering Using Genetic Algorithm Based on the Nearest Neighbour Heuristic. *International Journal of Computers and Applications*, **44**, 291-303. <https://doi.org/10.1080/1206212X.2020.1735035>
- [15] Kamar, M.E.Z.N., Esmailzadeh, A. and Heidari, M. (2022) A Survey on Deep Learning Techniques for Joint Named Entities and Relation Extraction. *Proceedings of 2022 IEEE World AI IoT Congress (AIIoT)*, Seattle, 6-9 June 2022, 218-224.
- [16] Liu, R. and Yu, H. (2022) Federated Graph Neural Networks: Overview, Techniques and Challenges. (Preprint)
- [17] Devlin, J., Chang, M.W., Lee, K., *et al.* (2018) Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. (Preprint)
- [18] Dosovitskiy, A., Beyer, L., Kolesnikov, A., *et al.* (2020) An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. (Preprint)
- [19] Abadani, N., Mozafari, J., Fatemi, A., *et al.* (2021) ParSQuAD: Machine Translated Squad Dataset for PERSIAN Question Answering. *Proceedings of 2021 7th International Conference on Web Research (ICWR)*, Tehran, 19-20 May 2021, 163-168. <https://doi.org/10.1109/ICWR51868.2021.9443126>
- [20] Seo, M., Kembhavi, A., Farhadi, A., *et al.* (2016) Bidirectional Attention Flow for Machine Comprehension. (Preprint)
- [21] Wang, W., Yang, N., Wei, F., *et al.* (2017) Gated Self-Matching Networks for Reading Comprehension and Question Answering. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, Vancouver, 30 July-4

August 2017, 189-198. <https://doi.org/10.18653/v1/P17-1018>

- [22] Yang, Z., Dai, Z., Yang, Y., *et al.* (2019) Xlnet: Generalized Autoregressive Pretraining for Language Understanding. *Proceedings of Annual Conference on Neural Information Processing Systems 2019 (NeurIPS 2019)*, Vancouver, 8-14 December 2019, 1-11.