DLINE JOURNALS

International Journal of
Computational Linguistics Research

Print ISSN: 0976-416X
Online ISSN: 0976-4178

IJCLR 2025: 16 (1)
https://doi.org/10.6025/ijclr/2025/16/1/10-17

Designing Neural Network-based Intelligent Robot English Translation System

Chunye Zhang, Tianyue Yu, Yingqi Gao
Harbin University of Science and Technology
150080, Harbin, Heilongjiang, China
zhangchunye@hrbust.edu.cn

ABSTRACT

Under the current wave of artificial intelligence, machine translation is a research direction in natural language processing, which has significant scientific and practical value. By analyzing the source language text in depth, we can utilize precise language features such as speech, vocabulary, intonation, etc., to transform it into understandable language results. Additionally, to address the machine translation task with a massive amount of language samples, the advantages of annotated data can be leveraged by combining the characteristics of neural networks to construct a more accurate translation model. Finally, by applying transfer learning, model overfitting can be avoided, greatly enhancing its generalization performance in less external input environments.

Keywords: Neural Network, Transfer Learning Technique, Intelligent Robot, English Translation, End-to-end Training

Received: 27 September 2024, Revised 22 December 2024, Accepted 29 December 2024

1. Introduction

In today's digitalized and globalized world, artificial intelligence and machine learning technologies are developing rapidly, and intelligent robots, as an essential part of them, have broad application prospects [1]. Intelligent robots must be able to communicate and interact with humans. English, as a universal language for global business, technology, and culture, makes English translation functionality one of the key components of intelligent robots. Traditional English translation systems mainly rely on rule-based or statistical methods. Still, these methods often need to thoroughly consider factors such as language ambiguity, contextual information, and cultural background, resulting in low translation accuracy and efficiency [2]. In recent years, with the advancement of deep learning technology, neural network-based translation systems have gradually become a

research focus, as they possess powerful semantic representation capabilities and good generalization performance, enabling them to handle complex language phenomena better. Neural networks are computational models that mimic the functioning of the human neural system, composed of multiple interconnected neurons [3]. The core of neural network algorithms is to optimize the weight parameters of the neural network through the backpropagation algorithm. The backpropagation algorithm is an iterative algorithm used for training neural networks, which continuously adjusts the weight parameters of the neural network by propagating errors from the output layer to the input layer to optimize the network's performance [4] gradually. Before using neural networks, all input information needs to be processed, and based on the processed information, accurate predictions of the output values of each neuron are made to ensure the final model is correct. Then, based on the error values, starting from the output layer, the errors of each neuron are calculated layer by layer until the input layer is reached. Finally, based on the error values and optimization algorithm, the neural network's weight parameters are updated to reduce errors gradually. The above steps are repeated until the network's performance reaches the expected level or maximum number of iterations.

This article further explores the design of a neural network-based intelligent robot English translation system. First, we elaborate on the overall architecture of the system and the functions and implementation methods of each component, including data collection, data preprocessing, neural network models, translation engines, and user interfaces. Then, we describe the algorithm design of the system, including end-to-end training methods, backpropagation algorithm, and stochastic gradient descent methods, and other optimization techniques. Next, we introduce the experimental part, including experimental settings, datasets, evaluation metrics, and analysis of experimental results. Finally, we summarize the research results and point out the future research directions and potential application prospects. Through the research in this article, we validate that the neural network-based intelligent robot English translation system has excellent performance in translation accuracy, speed, and user satisfaction. This system can effectively assist intelligent robots in English communication and interaction with humans and improve the application level of intelligent robots in fields such as education, healthcare, and tourism. At the same time, we will continue to explore the application of this technology in other natural language processing tasks to promote the further development of artificial intelligence and machine learning technologies.

2. Related Work

Neural Network-based Intelligent Robot English Translation System adopts deep learning technology and is trained on a large amount of bilingual data to achieve high-quality translation results. Among them, Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) are the most commonly used neural network structures [5]. RNN can handle sequential data and capture temporal information in the text, while CNN is suitable for processing structured data such as images and speech. In addition, the attention mechanism is also a widely used technique in neural machine translation in recent years, which can improve the accuracy and quality of translation. Previous research has also made extensive efforts in the design of intelligent robot English translation systems. Jiang et al. conducted a comprehensive review of the application of neural machine translation, discussing its basic principles, model architecture, training methods, evaluation metrics, and application scenarios. They also pointed out the shortcomings of current research and possible future research directions [6].

Lee et al. introduced the latest progress and trends in neural machine translation, including improvements in deep learning algorithms, optimization of model structures, and parallel training techniques. They discussed

the challenges and possible future development directions of neural machine translation [7]. Chen et al. proposed a parallel factorization method for parameter optimization in neural machine translation, which improves the efficiency and translation quality of the model by decomposing parameters into multiple subspaces and updating them with different learning rates [8]. Furthermore, Shen et al. proposed a Dynamic Time Warping (DTW) based neural machine translation method to handle translation problems in time series data[9]. By introducing the idea of time warping, the time series of the source language is mapped to the time series of the target language, which improves the translation accuracy and robustness of neural machine translation for time series data. Nagaraj et al. proposed an unsupervised domain adaptation method for the small-sample learning problem in neural machine translation, which improves the generalization ability and translation performance of neural machine translation systems with few labeled data by using unlabeled data for domain adaptation [10].

The Neural Network-based Intelligent Robot English Translation System applies deep learning and natural language processing technologies. It can achieve high-quality English translation effects through the learning and training of neural networks. This system has broad application prospects in business meetings, travel consultations, cross-language communication, education and training, and can provide people with more convenient and efficient translation services. With the continuous development of technology, intelligent robots will be applied and developed in more fields.

3. Construction of English Fusion Neural Network Transfer Learning Algorithm Translation Model

3.1 Transfer Learning Algorithm Combination Algorithm

In recent years, due to the outstanding achievements of Deep Convolutional Neural Networks (DCNN) in computer vision, using the features provided by CNN technology dramatically improves experimental efficiency and significantly increases recognition accuracy. Recent studies have found that features learned through DCNN are superior to manual interventions regarding recognition accuracy. DCNN, as a self-learning feature representation technology, has accuracy far surpassing traditional classic algorithms such as SIFT and HOG, and in practical applications, its accuracy also far exceeds traditional simulation techniques. Its generalization performance is also excellent, as it can quickly transform into other forms from any perspective, thus achieving the best results in practical scenarios. By adopting DCNN technology, we can greatly improve the recognition efficiency of fine-grained images, achieving accuracy and generalization far beyond manual design. For deep convolutional neural networks, information preprocessing is crucial. It not only allows information to achieve a standardized normal distribution but also controls the size of weights through multi-level network structures to ensure the accuracy of the final results. We introduce Batch Normalization (BN) technology to effectively suppress the deviation of internal covariances, transforming the distribution of hidden unit values into a capturable form, thereby enhancing the accuracy and reliability of the model. It also effectively adjusts and optimizes parameters, speeding up the model's training progress. By introducing a new weight decay term to suppress overfitting of the loss function, this measure can be implemented through formula (1).

$$g_{i,j,f}^{1} = \mathrm{Re}\,LU\left(b_1 + \sum_{a=1}^{m}\sum_{b=1}^{n} g_{i,j,f} + b \cdot t_{a,b,c}^{1}\right) \tag{1}$$

3.2 Neural Probability Algorithm

Although N-gram language models are relatively simple, their training results often lack sufficient N-gram

words, leading to insufficient accuracy and data sparsity. Therefore, to improve the model's accuracy, some smoothing algorithms must be adopted to enhance their performance. Some common algorithms, such as additive smoothing, Kneser-Ney, Katz, and Jelinek-Mercer, can handle complex contextual relationships. However, as the context length increases, the utilization rate of the n-gram grammar will sharply rise, limiting the model's ability to capture complex contextual relationships effectively. This is the biggest drawback of N-gram models. Therefore, the idea of applying neural networks to language models is proposed, and the exponential growth of parameters is overcome by sharing parameters between similar data. Studying the probability function of language can help us better understand and grasp the numbers. However, with the development of digital technology, this method becomes increasingly complex, especially when we need to deal with the distribution of a large number of discrete random variables and discrete distribution of digital data analysis. Adopting a new approach, using different combinations of letters, can also accurately identify a group of letters with similar meanings. Moreover, it can better understand the meaning of a group of letters based on the continuous changes of letter combinations, thus better predicting their meanings. Statistical language models can be represented as conditional probability multiplication, as shown in Formula 2. In this experiment, we found that the neural network structure not only includes a feature vector mapping layer but also two hidden layers: one is the C-layer, which is responsible for sharing language features, and the other is the hyperbolic tangent layer, which converts input information into a hyperbolic tangent to process the output information. In fact, the neural network also has a set of algorithms to ensure that its results are positive, and the calculation formula for each result is shown in Formula 3.

$$\hat{P}\left(\omega_1^T\right) = \prod_1^T \hat{P}\left(\frac{\omega_t}{\omega_1^{t-1}}\right) \tag{2}$$

$$\hat{P}\left(\omega_t, \omega_1^{t-1}, \ldots, \omega_{t-n+1}\right) = \frac{e^{y_{wt}}}{\sum e_i^{y_i}} \tag{3}$$

3.3 Encoder-Decoder Framework

When the sentence we need to translate is particularly long, the information vector may not be able to thoroughly represent all the information of our source input sentence, leading to information loss. One solution is to extract all the hidden outputs of the RNN encoding process into our decoder, which would be large and not concentrated. Therefore, an attention mechanism is introduced to extract all the encoded information. Figure 1 shows the encoder-decoder framework with an attention mechanism. The many-to-many structure is used for machine translation, and this structure is also used for tasks like chat dialogues. The encoder-decoder framework has many applications. This framework typically refers to encoding an input into a message and decoding it to obtain an output; it is often used in image classification schemes. Firstly, the context length is precisely controlled by the stacking of convolutions. Secondly, convolutions allow parallel computation, as they are independent of the previous time step's state and allow each element in the sequence to be processed in parallel. Thirdly, the training complexity is reduced as the number of convolution kernels and non-linear computations for each word in the input convolution network is fixed. Many frameworks typically refer to encoding one input to generate multiple outputs; this framework is often used to input an image and generate multiple descriptions of that image. The many-to-one framework typically refers to accepting many inputs and ultimately producing only one output,

and it is often used in text sentiment classification, where multiple words in the input sentence input the sentiment of the sentence as output. Another many-to-many framework is the synchronous many-to-many framework, where each time step of the input has a corresponding output, and the output at this moment is fed back to the next moment as input. This can be used for character prediction, video tagging, video classification, etc.

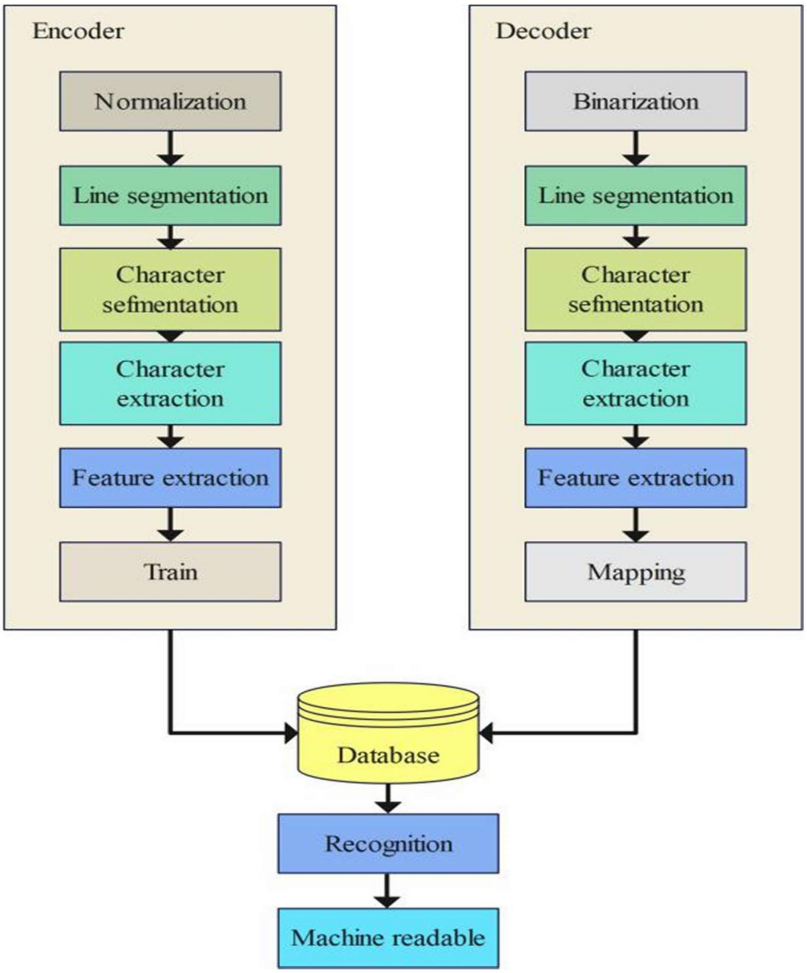

Figure 1. Encoder-Decoder Framework

4. Experimental Design and Analysis

4.1 Experimental Design

By using convolutional operations, complex information can be transformed into forms that can be processed nonlinearly in order to recognize better and describe various aspects of images and, thus, better grasp the information within them. Moreover, as each feedback is unique, it can better capture and interpret the semantic changes in the image. DCNN was developed to address complex image problems, and it can divide the model based on three critical dimensions (height, width, and depth) to achieve information interaction among multiple models. In contrast, CNN can achieve more efficient self-learning, greatly reducing the cost of human

intervention. By adopting DCNN technology, we can apply the original techniques to the processing of smaller-sized speech and images, thereby handling complex information more effectively. The advantage of this technology is that it can use the original methods to the processing of smaller-sized information, thereby handling complex information more effectively and better adapting to various needs. Through the massive image resources of ImageNet, we can effectively train DCNN to obtain its effective weights and feature representation capabilities. Then, we can apply this technology to analysing smaller images, thereby capturing and processing information more effectively and achieving better image classification and processing. English is a low-resource language, with few English corpora, especially in certain specific domains. Therefore, training a neural network with few English corpora will not work well. The solution proposed in this study is first to train an English-Chinese neural machine translation model, then use transfer learning techniques to transfer it to the original model, and finally obtain an English-Chinese neural machine translation model. In real machine translation tasks, the performance of a translation model trained on one language using a parallel corpus dataset for another language will significantly degrade because the distributions of the test and training sets will differ. Domain adaptation is a type of transfer learning that enhances the performance of the target domain model by using data samples from the information-rich source domain. Neural machine translation maps the translation language to input sequence information and then generates the target language. During this process, all the information and features of the translation language are stored in the background vector. The background vector can convey semantic information and integrate more language knowledge to improve encoding efficiency and translation performance. Information is stored in the background vector as prior knowledge during encoding, collectively achieving translation training.

4.2 Experimental Results

The transfer efficiency is affected by the variation between the source domain and the target domain. If the variation is too large, it may lead to negative transfer effects. As shown in Figure 2, the original source domain information can assist the simulation of the target domain. Still, it can also influence the final results, leading to a decrease in simulation efficiency. However, despite this, some commonalities exist in the core structures of text convolution and SAR convolution. By comparing the visibility of the convolution kernels in AlexNet and Mnet2, we find that their core appearances are roughly consistent, allowing them to better support data transformation. The alignment matrix of the source sequence and the target sequence shows the important distribution of each word in the source sequence that needs to be translated when translating a current word. When a model's pre-processing performance exceeds its expected range, it may experience overfitting. In this case, although its pre-processing performance can reach a reasonable level, its generalization ability is not ideal, as it cannot effectively eliminate errors in pre-processing. When using transfer learning techniques for model training, if overfitting occurs, it will lead to a significant decrease in the convergence rate of *Train_loss*, resulting in substantial fluctuations in the model's accuracy *Test_accuracy*. Based on Figure 3, to effectively avoid overfitting, the following techniques are adopted: reducing the model's size, interrupting training promptly, and applying regularization; L1 regularization can make the weight matrix appear denser, while L2 regularization can minimize it. By using two different regularization operators, one can effectively reduce the size of the grid, preventing excessive deformation, and the other can effectively suppress sampling bias by using a large amount of training data, further inhibiting overfitting. In addition to these commonly used methods, network pruning can also alleviate the overfitting problem. Early convolutional neural network technology has been proven to have good

network pruning capabilities. It can not only reduce the complexity of the model and avoid overfitting but also eliminate redundant parameters, thereby improving the overall system's generalization performance. By combining two different datasets, we can create a dual-stream CNN. This CNN can constrain the difference between the two datasets through the loss function, making their outputs as close as possible. This way, both datasets' models can accept and effectively perform data transformation, achieving efficient data propagation. Through deep learning of the MSTAR10Class 22-type data, we found that the model pruning technique can significantly improve the accuracy of transfer learning and avoid overfitting. Moreover, our experiments also indicate that this technique can be applied to complex heterogeneous network environments.

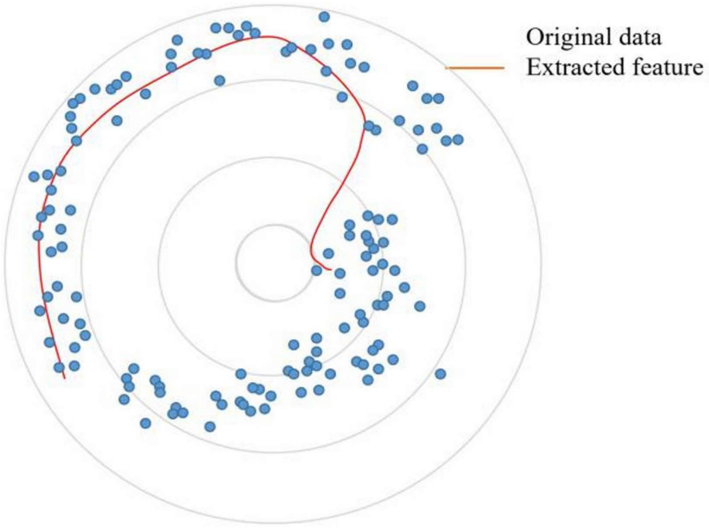

Figure 2. Comparison of distribution of original data and extracted features

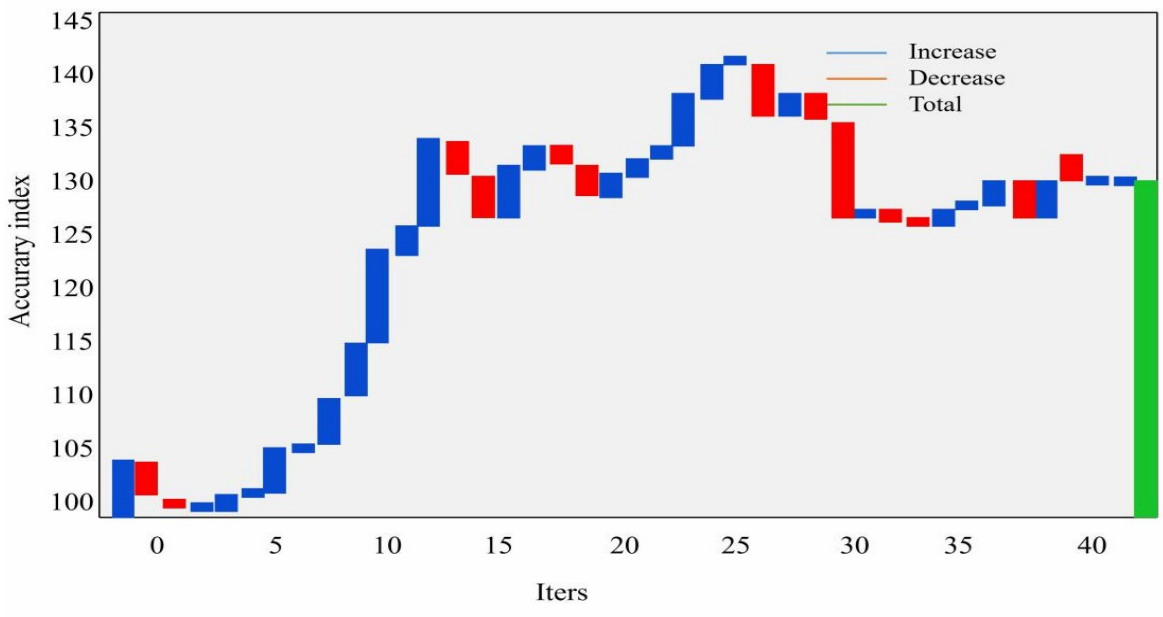

Figure 3. Overfitting of Transfer Learning Neural Network Model

5. Conclusion

The research work in this paper mainly focuses on English-to-Chinese neural machine translation, covering aspects such as bilingual corpora, language features, neural networks, and transfer learning. Various factors influence English and possess many unique language features. Firstly, relevant research has been conducted on the specific language features of English, mainly focusing on lexical and syntactic features. Studying English's lexical and syntactic features is essential to improve the performance of neural machine translation models. Without large-scale parallel corpora, machine translation tasks can be transformed into unsupervised tasks by pretraining the language model using large-scale monolingual corpora to obtain the overall understanding. Unsupervised machine translation is suitable for improving machine English translation under resource-limited conditions.

References

[1] Jiao, L. (2023). A novel deep nearest neighbor neural network for few-shot remote sensing image scene classification. *Remote Sensing, 15*(1), Article 15.

[2] Seifi, P., Sani, S. K. H. (2023). Barrier Lyapunov functions-based adaptive neural tracking control for non-strict feedback stochastic nonlinear systems with full-state constraints: A command filter approach. *Mathematical Control and Related Fields, 13*(3), 988-1007.

[3] Sadek, E. T., Seada, N. A., Ghoniemy, S. (2023). Neural network-based method for early diagnosis of autism spectral disorder head-banging behavior from recorded videos. *International Journal of Pattern Recognition and Artificial Intelligence, 37*(05), 2356003.

[4] Wang, X. (2021). Image recognition of English vocabulary translation based on FPGA high-performance algorithm. *Microprocessors and Microsystems, 80*, 103542.

[5] Liu, T. (2021). Prediction of tomato yield in Chinese-style solar greenhouses based on wavelet neural networks and genetic algorithms. *Information, 12*(1), Article 12.

[6] Jiang, Z., Li, B., Tran, T. N. H. T., et al. (2022). Fluo-Fluo translation based on deep learning. *Chinese Optics Letters, 20*(3), 031701.

[7] Lee, Y. H., Shin, J. H., Kim, Y. K. (2021). Simultaneous neural machine translation with a reinforced attention mechanism. *ETRI Journal, 43*(5), 1-11.

[8] Chen, X. (2021). Simulation of English speech emotion recognition based on transfer learning and CNN neural network. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 40*(2), 1-12.

[9] Shen, X., Qin, R. (2021). Searching and learning English translation long text information based on heterogeneous multiprocessors and data mining. *Microprocessors and Microsystems, 82*(2), 103895.

[10] Nagaraj, P. K., Ravikumar, K. S., Kasyap, M. S., et al. (2021). Kannada to English machine translation using deep neural network. *Ingénierie des Systèmes d'Information, 26*(1), 123-127.