

¹Nan Zhang²Xiaozhou He*

Corpus-based Research on English-Chinese Translation Teaching Combining Vocabulary Learning and Practice



Abstract: - As the global economy is developing, real time translation among major influential languages such as English, Chinese has become a necessity. Translation from English to Chinese is difficult since Chinese has a different grammar and ambiguous word boundaries. Furthermore, there are issues with the current Chinese-English machine translation, including difficult-to-understand extended sentences and inaccurate word translation. To address this drawback, the methodology for more precisely categorizing vocabulary in both English and Chinese is described in this research. The data are collected from UM-Corpus news dataset. Then, using SEGBOT: Neural Text Segmentation method, English words are segmented from the sentences available on the dataset. Afterward, the data are fed to pre-processing. In pre-processing segment; words missing in translation from the news dataset are eliminated and enhances the input data utilizing Iterated Square-Root Cubature Kalman Filter. The outcome from the pre-processing data is transferred to the Pyramidal Convolution Shuffle Attention Neural Network (PCSANN). The word order, grammatical structure, consistent style and smooth flow for English to Chinese translation are successfully classified by using PCSANN. The Waterwheel Plant Algorithm (WWPA) is used to optimize the weight parameter of PCSANN. The proposed PCSANN-WWPA is applied in python working platform. Performance metrics, like accuracy, precision, F1-score, and recall are examined to compute proposed method. The gained results of the proposed PCSANN-WWPA method attains higher accuracy of 16.65%, 18.85%, and 17.89%, higher sensitivity of 16.34%, 12.23%, and 18.54% and higher precision of 14.89%, 16.89%, and 18.23%. The proposed ECTT-PCSANN-WWPA method is compared with the existing methods such as ECTT-RNN, ECTT-DNN, and ECTT-MLPNN models respectively.

Keywords: Waterwheel Plant Algorithm, Iterated Square-Root Cubature Kalman Filter, Pyramidal Convolution Shuffle Attention Neural Network, English-Chinese Translation, Neural Network.

I. INTRODUCTION

The two most significant languages in the global economy today are Chinese and English. Given China's increasing global importance, translating between English and Chinese is crucial for both understanding China and serving as a bridge between the two countries. Researchers have shown a great deal of interest in machine translation as a practical way to overcome the obstacles to cross-language communication [1, 2]. One of the primary uses of natural language processing (NLP) is machine translation. This kind of study on applied technology [3]. Big data is the foundation for the majority of machine translation technology used today. Training on a massive amount of data is required to get improved results [4]. The notable developments in deep learning research have led to major breakthroughs in neural machine translation (NMT) based on recurrent neural networks [5]. Significant advancements in recurrent neural network-based neural machine translation (NMT) have been made possible by the noteworthy advances in deep learning research [6]. With limited funding, this effort can offer fresh perspectives to shape the course of neural machine translation research. In conjunction with theoretical and empirical support for translation technology, engineering applications, and mutual translation training for low-resource language speakers across different regions [7, 8].

Statistical machine translation, the fundamental technology behind the translation platforms of major global corporations like Google, Baidu, and Sogou [9, 10], still has a lot of unresolved issues, including linear inseparability, data sparseness, and imprecise semantic representation [11]. In language pairs, Sometimes there are only a few hundred to thousands of sentences that are parallel [12]. It is highly challenging to train an effective machine translation system in the absence of data since SMT and NMT rely heavily on data [13]. NMT and SMT are both very dependent on data. Currently, automatic big data training is the source of automated translation knowledge. The efficacy of automated translation is impacted by the fact that the knowledge gained from this training is still very different from that of interpreters with extensive training [14]. Compared to statistical machine translation, NMT is more sensitive to sentence length [15], and the end-to-end implementation process does not explicitly leverage linguistic knowledge to enhance translation efficiency [16].

¹School of Foreign Studies, Suqian University, Suqian 223800, Jiangsu, China. zhangnan@squ.edu.cn

²Public Course Teaching Department, AnHui Business Vocational College, Hefei, Anhui, 231131, China. hezhouT@163.com

Copyright © JES 2024 on-line: journal.esrgroups.org

To improve the model's training effect, a parallel corpus of English and Chinese will be built, and the English and Chinese corpus data with considerable data value will be taken from scholarly publications [17, 18]. The optimal window size is ascertained by comparison when developing and testing the multilayer perceptron network structure for text categorization with windows of different sizes [19, 20].

Testing various hidden layer network layers also determines the amount of network layers in this model. After receiving English text as input, the PCSANN-WWPA algorithm converts it to Chinese output using Google's Transformer-based encoder-decoder model. Considering this translation system is platform-portable, it can be used in small- to medium-sized business settings and for educational purposes. Many factors are taken into consideration while analyzing the impact of English-Chinese translation, including recall, accuracy, precision, and f1-score. significant input to the work that is compiled here;

- At first, the data is gathered via the news dataset of UM-corpus.
- Using Iterated Square-Root Cubature Kalman Filtering to eliminate the words missing in translation from UM-corpus news dataset in the pre-processing segment.
- The pre-processed data are fed into the Pyramidal Convolution Shuffle Attention Neural Network in order to effectively categorize English-Chinese translation of the UM-corpus news dataset into word order, grammatical structure, consistent style and smooth flow.
- The proposed PCSANN-WWPA approach is used, and performance metrics such as F1-score, recall, accuracy, and precision are examined.

This manuscript's remaining sections are organized as follows: Part 2 looks at a review of the literature; Part 3 describes the recommended approach; Part 4 shows the findings and discussion; and Part 5 wraps up.

II. METHODOLOGY

In the literature on teaching English to Chinese translation, a number of study projects were suggested; a few recent works are reviewed here,

Xu [21] have suggested that a transfer learning-based English-Chinese machine translation study methodology. It begins by outlining the principles of transfer learning, neural machine translation and related fields of study. The transformer neural machine translation model framework was selected after research on neural machine translation and other models' advantages and disadvantages were discussed. Thirty million parallel Chinese-English corpora, one hundred thousand low-resource Chinese-English corpora, and one hundred thousand. The transformer machine translation architecture was pre-trained using parallel corpora of Tibetan and Chinese. The transformer uniform distribution was used to initialize the model parameters, the decoders consisted of six identical hidden layers, and Adam was used as the optimizer during the model training process. The pre-trained model's parameters were used to the resource-efficient machine translation models for Tibetan-Chinese and Chinese-English to fulfill the objective of knowledge transfer.

Sun [22] have suggested that in the Chinese machine translation training era of information, people can build effective communication over the Internet. Nowadays, one of the most important tools for resolving communication issues is mechanical translation. There were still barriers to cross-linguistic communication, though. To address this difficulty in the field of marine science and technology, this work integrates the most recent advancements in neural network technology to the English-Chinese bidirectional machine translation model. Utilizing deep learning technology, compile abstracts and partially written articles about marine science and technology in both Chinese and English. Using these themes as keywords, create an expert corpus about marine science and technology in both Chinese and English. We improved the translation in both directions paradigm between Chinese and English by including local weight sharing between the encoders for Chinese and English. The effectiveness of the translation model was evaluated using the BLEU parameters after the output of the Chinese and English encoder sub layers was fused together to represent the output of their respective encoders.

Xue and Wang [23] Have suggested that an MLP structure made of two layers of feed forward neural networks was created to accommodate the features of the English-Chinese translation corpus. Additionally, a text classifier utilizing multilayer perceptrons was offered. Training samples were selected based on text heights, mainly less than 60 pixels. A mathematical morphological smoothing method was applied to select the optimal 13x13 sliding window by comparison. After being put into practice to teach English-Chinese translation in colleges and universities, the developed MLP-based information-based teaching model was primarily assessed in terms of student evaluation and instructional efficacy. Students' passion for participating in English-

Chinese translation learning can be sparked by MLP-based informatization instruction, which can also reinforce the direction of translation techniques and raise students' translation proficiency.

Zhao and Jin [24] have suggested that traditional translation techniques have gradually been replaced by English-Chinese translation models based on neural networks. The "encoder-attention-decoder" The neural translation model's primary tool for simulating the entire translation process is structure. Comprehending grammar was also essential to the translation process because it reduces grammatical errors and aids in appropriately depicting word sequences. The following two studies will be conducted using the attention mechanisms and grammatical knowledge that was the subject of this article. First, in light of the current neural network structure, A long short-term memory (LSTM) network translation model was introduced as a kind of embedded attention to build a translation model influenced by long-distance information lost during transmission. which led to issues regarding the translation effect, which was not ideal. Secondly, given that translation models do not incorporate grammatical prior knowledge, a technique to incorporate grammatical information as previous knowledge into translation models is needed.

Li [25] have suggested a unique machine translation model that integrates transfer learning with translation expertise, based on bilingual translation between Chinese and English. It is composed of a recurrent neural network-based translation quality evaluation model and a self-focused network-based model. The current Chinese-English machine translation suffers from issues like imprecise word translation and challenging sentence translation.

Ruan [26] have suggested that worldwide language intercommunication and translation have emerged as essential requirements for smooth worldwide human communication. The development of machine translation from academic research to commercial applications was made possible by the advancement of computer technology. Furthermore, deep learning is a relatively new and well-liked subfield of machine learning that has excelled in areas of study including organic language interpretation. By examining the intelligent recognition of English-Chinese machine translation models, this work enhanced the effectiveness of deep learning network-based machine translation. Finding answers was this study's main objective for machine translation's out-of-vocabulary (OOV) problems involving uncommon and unregistered words. In addition, It brought together byte pair encoding (BPE), word sequence segmentation techniques based on subwords, and stemming technology. In order to focus on the relationship between the context vectors at various points in the decoder, this work used a two-layer computational structure. It was inspired by the conventional attention calculation procedure. Using the neural machine translation model developed by GNMT, this work carried out an experimental investigation on three different scale datasets using the enhanced methodologies mentioned above.

Zhao et al. [27] have suggested that there is a lot of use for machine translation automation in the fields of international trade, healthcare, tourism, education, and text digitization.. Because Chinese has a different grammar and unclear word boundaries from word-based languages like English, translating words into Chinese can be challenging. Utilizing a corpus with a specific domain, A machine translation system with deep learning with GPU support has been built by this research. Using an attention mechanism inspired by Google's Transformer and an encoder-decoder paradigm, After taking an English input, our algorithm generates a text in Chinese. The model was trained with bilingual English-Chinese text phrases from the News section of the UM-Corpus using an Adam optimizer and a straightforward self-designed entropy loss function.

A. Motivation

The overall analysis of the most current research demonstrates that, English-chinese translation teaching is an important means for seamless communication. The lack of data makes training a machine translation system difficult. Many researchers deal that problem with the different technologies like recurrent neural network, deep neural network, multiperceptron neural network and Byte pair encoding (BPE) algorithm. By creating a Chinese-English corpus, The deep neural algorithm-based translation model amplifies the effect of model training. Neural network techniques are applied to nonlinear problem classification. The BPE algorithm maintains a balance between translation efficiency and vocabulary size. Tokenization is applied to the original dataset using the BPE technique. During the inference stage, the optimal response is located using the beam search technique. It broadens the area of solutions and decreases complexity in comparison to the greed search; on the other hand, the thorough investigation increases complexity. Size of beams is a critical component affecting the efficacy and efficiency of the beam search process. To convert English text to Chinese output, An attention mechanism modeled after Google's Transformer is used with an encoder-decoder model in the deep

learning approach. There aren't many approach-based publications in the literature that address this issue; these shortcomings and issues are what motivated the research work.

III. PROPOSED METHODOLOGY

This section discusses research on teaching English to Chinese translation by integrating vocabulary acquisition with practice using a hybrid technique. Based on the corpus data, this paper offers a novel method for teaching English to Chinese translation. Fig. 1 displays the proposed methodology's block diagram. There are four steps in it: gathering data, segmenting it, preprocessing it, and classifying it.

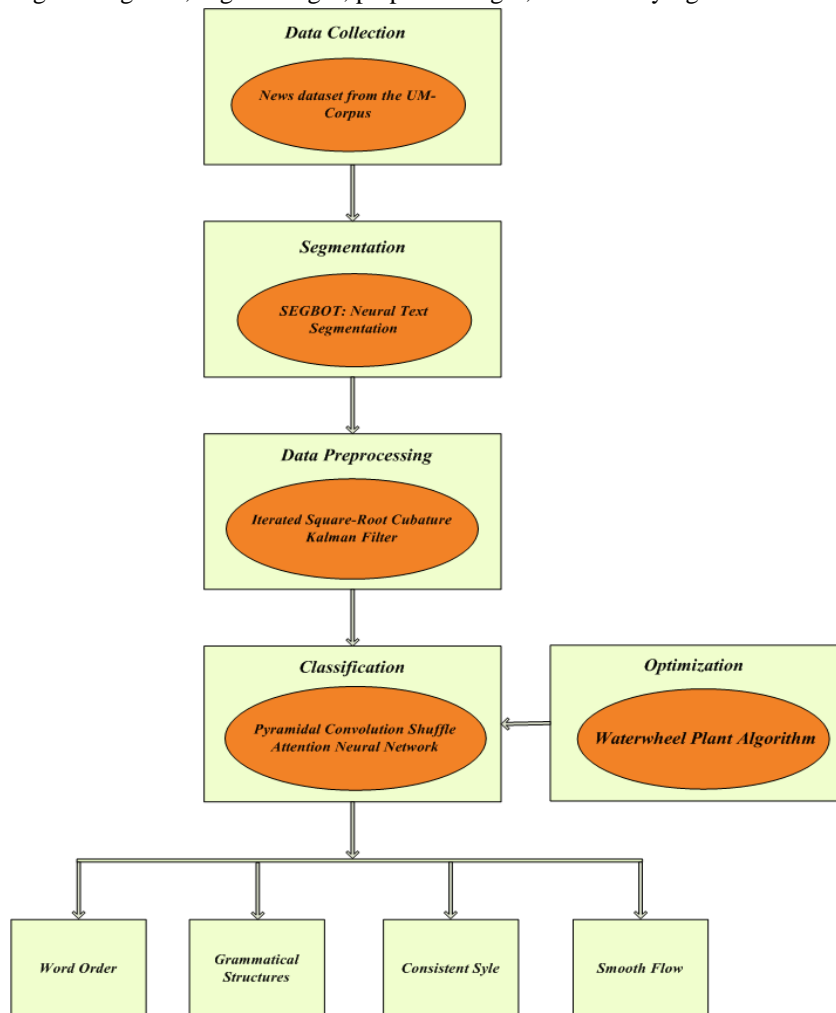


Fig 1: Block Diagram of Proposed Methodology

A. Data Acquisition

The information came from the large parallel English and Chinese UM-Corpus news dataset [28]. There are two million English-Chinese text corpora available, spanning eight text domains and covering a wide range of topics and text genres such theses, spoken English, education, laws, news, science, microblogs, and subtitles.

The news subset, which contained 252K phrases, 10,635K words in Chinese and 5672K words in English was used to train our models. From the training samples, 25,278 validation pairings, 50,556 test pairs, and 176,943 training couplets were then produced.

B. Segmentation using SEGBOT: Neural Text Segmentation

The process of dissecting a written language string into its individual words is known as text segmentation. To solve the text segmentation issue of varying output vocabularies and sparse boundary tags, SEGBOT: Neural Text Segmentation is used [29]. SEGBOT has the capacity to segment text at different granularities. Boundary pointer, boundary decoder, and context encoder make up SEGBOT's three parts.

1) Context Encoder

The input sequence should be encoded. $Y = (y_1, y_2, \dots, y_N)$ using an RNN. Input sequences are encoded using GRU, a less expensive computational method that is comparable to LSTM.

The GRU activations at time step n are computed as follows:

$$k_n = \sigma(W_z y_n + R_z x_{n-1} + c_z) \tag{1}$$

$$l_n = \sigma(W_i y_n + R_i x_{n-1} + c_i) \tag{2}$$

$$m_n = \tanh(W_j y_n + R_j (l_n \circ x_{n-1}) + c_j) \tag{3}$$

$$p_n = k_n \circ x_{n-1} + (1 - k_n) \circ m_n \tag{4}$$

where the element-wise multiplication is \circ , the update gate vector is k_n , the reset gate vector is l_n , the new gate vector is m_n , the sigmoid function is $\sigma()$, the hyperbolic tangent function is $\tanh()$, and the hidden state at time step is x_n . These are the parameters of the encoder to learn n . W, R, c

2) Boundary Decoder

Given that RNN-based models can handle varying sequence lengths, Given that the number of boundaries in the output varies based on the input, using them to decode the output makes logical. At every time step, the decoder's concealed state is determined in the manner described below:

$$f_m = GRU(y_m, \theta) \tag{5}$$

Where, θ are the decoder's hidden layer parameters.

3) Boundary Pointer

At each step, the output layer of our decoder determines a probable segment boundary by calculating its distribution across all possible places in the input sequence. The distribution across all possible input sequence points for input symbol decoding V_m is computed using an attention method.

$$u_a^m = v^T \tanh(W_1 x_a + W_2 f_m) \tag{6}$$

$$p(x_m | y_m) = \text{soft max}(u^m) \tag{7}$$

where, given the start unit U_m , normalizes u_a^m , to indicate the likelihood that the unit U_a is a boundary, and $a \in (m, M)$ denotes a potential location in the input sequence. When considering the start unit, U_a , u_a^m , normalizes to show the probability that the unit U_m , is a boundary, and $a \in (m, M)$ denotes a possible position in the input sequence.

The segmenter divides various affixes, like possessives, and splits punctuation for English. Whitespace is treated like any other symbol. As a result, the text can be de-tokenized without any confusion. We can efficiently do word segmentation with this technique.

C. Preprocessing using Iterated Square-Root Cubature Kalman Filter (ISCKF)

Given that the SCKF employs the Newton-Gauss iterative strategy, it is expected that the ISCKF methodology will perform better in terms of numerical stability and accuracy. [30].

The forecasting tool projects the state vector by using the expected state vector from the previous iteration. Assuming a first-order $1 \times 10^{-3} \times I$ state covariance matrix and a flat initial voltage profile.

For the given state vector, cubature points are obtained by:

$$A_{i,j-1|x-1} = Q_{x-1|x-1} \xi + \hat{f}_{x-1|x-1} \quad i = 1, 2, \dots, m \tag{8}$$

Where, $m = 2n$ and $j = 1$

$$Q_{0|0} = \text{sqr}(O_0) \tag{9}$$

Utilizing $f()$, determine the cubature points ($i = 1, 2, \dots, m$) in

$$A_{i,j-1|x-1} = f(A_{i,j-1|x-1}) \tag{10}$$

The projected state vector and the observations for time instant k are used to calculate the filtered state vector and Kalman gain. To provide an even stronger filtering effect, Recursively, this is carried out for iterations $p = 0, 1, \dots, N_{iter}$ where $N_{iter} = 2n$ in our case.

Analyze the cubature points for the filtering stage ($i = 1, 2, \dots, m$),.

$$A^{(z)}_{i,j-1|x-1} = \hat{Q}^{(z)}_{x-1|x-1} \xi_i + \hat{f}^{(p)}_{x|x-1} \tag{11}$$

In the case of the filtering phase, number of internal iterations $j = 0, 1, \dots, N_{iter}$ for $j = 0$, $\hat{Q}^{(0)}_{x|x-1} = Q_{x|x-1}$ and $\hat{f}^{(0)}_{x|x-1} = \hat{f}_{x|x-1}$

Assess the propagation of cubature points.

$$P^{(z)}_{i,x|x-1} = h\left(A^{(z)}_{i,x|x-1}\right) \tag{12}$$

The data is preprocessed and the words missing in translation from the news dataset were eliminated and the preprocessed data is fed to classification.

D. Classification by Pyramidal Convolution Shuffle Attention Neural Network (PCSANN)

In order to capture more complex characteristics without raising the computing cost, a pyramidal convolution combines a number of convolution kernels at different scales with varying spatial resolution and depth [31]. Furthermore, the convergence layer can identify complex features since the pyramidal convolution is able to recognize the relationships between spatial features at various levels. The features are split into groups that are expressed in equation (13).

$$Y = [Y_1, Y_2, \dots, Y_N] \tag{13}$$

Where, Y represents the feature map, N represents the number of groups. Then the function efficiency is enhanced by the use of the linear function. Finally, global information is introduced through embedded with the original assets, after its activation sigmoid function to obtain a class representation, it is given in equation (14).

$$Y_{a1} = \gamma(h_w(q))Y_{a1} = \gamma(p_1q_1 + k_1)Y_{a1} \tag{14}$$

With γ is c representing the linear function h_w and standing for the average pooled function, q_1 implies the average pooled function, and p_1, k_1 are obtained from network training. Spatial perception can be preserved attention improvement. The spatial attention function is given in equation (15).

$$Y_{a2} = \gamma(p_2 \cdot HR(Y_{a2}) + k_2) - Y_{a2} \tag{15}$$

where q_2 is the normalized feature and HR stands for the group norm normalization function. The two branches are spliced and aggregated, and then it is given in equation (16).

$$Y_{a2} = [Y_{a1}, Y_{a2}] \times P \times T \tag{16}$$

Where, P, T are denoted as the sub features. Then every sub functions are grouped. Finally, channel grouping procedure is carried out. The completely connected layer's output vector moves forward through a sigmoid layer, this is expressed in equation (17).

$$\tilde{l}(w|J) = 1 / (1 + \exp(-lh(w|J))) \tag{17}$$

Where, J is the input data, $\tilde{l}(w|J)$ is the probability score, the global branching is given in equation (18).

$$P(C) = \frac{-1}{W} \sum_{w=1}^W p_w \log(\tilde{l}(w|J)) + (1 - p_c) \log(1 - \tilde{l}(w|J)) \tag{18}$$

Where, p_w represents the true label, w and W represents the total number of teaching categories.

PCSANN accurately classifies the English translation Typically, PCSANN lacks the optimization methods needed to choose the best variables to validate a precise detection .Therefore, in order for the optimization method to optimize PCSANN, the weight parameter γ is crucial.

E. Optimization Using Waterwheel Plant Algorithm (WWPA)

Waterwheel plants are known for their broad petioles, which carry their small, translucent flytrap-like traps [32]. In order to guard against damage or accidental triggers from other water plants, a ring of hair-like bristles encircles the trap. Many hook-like teeth are coated around the outside edges of the trap, and while the victim is encircled by the trap,, the teeth interlock just like those on flytraps. To help with meat digestion, predators have

glands that secrete acid and trigger hairs. The victim is drawn into the trap by the teeth that lock together and the mucus sealant that pins them to the floor at the hinge.

Step 1: Initialization

Set the input parameters to their initial values. The PCSANN weight parameters, represented by the letter γ , are the input parameters in this case.

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,m} \end{bmatrix} \tag{19}$$

$$X_{i,j} = yb_j + a_{i,j} \cdot (zb_j - yb_j), i = 1, 2, \dots, N, j = 1, 2, \dots, m \tag{20}$$

where the numbers m and N stand for the quantity of waterwheels and other factors, correspondingly; The waterwheel locations' population matrix is represented by X . X represents the positions of the waterwheels' population matrix; yb_j and zb_j are the j-th problem variable's lower and upper bounds; and $a_{i,j}$ is an arbitrary number that falls between 0 and 1.

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_N \end{bmatrix} = \begin{bmatrix} A(B_1) \\ \vdots \\ A(B_i) \\ \vdots \\ A(B_N) \end{bmatrix} \tag{21}$$

The expected value for the i -th waterwheel is represented by A_i , and vector A contains all of the objective function values.

Step 2: Random Generation

A matrix's input parameter is created at random.

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix} \tag{22}$$

Where, X is the population of waterwheel location, The quantity of variables is d , while the number of waterwheels is n .

Step 3: Fitness Calculation

To calculate the fitness value by F

$$Fitness\ function = Optimizing(\gamma) \tag{23}$$

Step 4: Exploration Phase

With their excellent sense of smell, waterwheels can identify the exact location of pests, making them formidable predators. Any bug inside the waterwheel's attack range will be targeted. Once it has found its victim, it attacks and keeps searching. WWPA simulates this waterwheel action in order to mimic the initial step of its population updating procedure. WWPA may be better able to explore the search space, find the ideal region, and escape from local optima if it replicates the insect's attack on the waterwheel, leading to notable variations in the waterwheel's position.

$$X = i_1 \cdot (U(j) + 2V) \tag{24}$$

$$U(j+1) = U(j) + X \cdot (2V + i_2) \tag{25}$$

where the random variables i_1 and i_2 have values in the interval $[0, 2]$ to $[0, 1]$. Furthermore, X is the circle's circumference, which the waterwheel plant will utilize to look for potential places, and V is a variable with exponential values between 0 and 1. If the solution doesn't change after three rounds, Equation (25) can be used to shift the placement of the waterwheel.

$$U(j+1) = \text{Gaussian}(\mu\rho, \sigma) + i_1 \left(\frac{U(j) + 2V}{X} \right) \quad (26)$$

Step 5: Exploitation Phase

A waterwheel attracts an insect, which is then moved through a feeding tube. The second phase of the WWPA population update is informed by this mimic waterwheel activity. Every waterwheel in the population is determined by the WWPA's designers to be a new, random site that is a "excellent position for devouring insects" in order to mimic the natural behavior of waterwheels. The following equations demonstrate that the waterwheel is shifted to the new location if the objective function value is higher there than it was at the previous site.

$$X = i_3 \cdot (VU_{best}(j) + i_3 U(j)) \quad (27)$$

$$U(j+1) = U(j) + VX \quad (28)$$

where U_{best} is the optimal solution, i_3 is a random variable having values between 0 and 2. and $U(j)$ represents the current solution at iteration j .

$$U(j+1) = (i_1 + V) \sin\left(\frac{X}{Y} \theta\right) \quad (29)$$

In this case, the random variables X and X have values between [-5, 5]. Furthermore, using equation (30), the value of V decreases exponentially.

$$V = \left(1 + \frac{2 * j^2}{T_{max}} + X \right) \quad (30)$$

Step 6: Update the Best Solution

The process is finished if the best result is achieved.

Step 7: Termination

If the solution is the best, the procedure will terminate; if not, it will loop back to the step 3 fitness calculation and continue processing the subsequent levels until a solution is found.

IV. RESULT AND DISCUSSION

The test outcome of Corpus-based Research on English-Chinese Translation Teaching using PCSANN-WWPA method is discussed in this session. Using the Python working platform, the proposed method is simulated under a number of performance criteria. Outcome of ECTT-PCSANN-WWPA is analyzed with existing methods such as ECTT-RNN, ECTT-DNN and ECTT-MLPNN.

A. Performance Measures

In order to determine the optimal classifier, this is an important task. Performance indicators including recall, F1-Score, precision, and v are examined in order to assess the performance. The performance measures will be scaled using the confusion matrix. it is decided. The True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP) values are required in order to scale the matrix of confusion.

1) Accuracy

It is the ratio of count of exact prediction and total number of forecasts for a given dataset. It is measured through equation (31).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (31)$$

2) F1-Score

In order to assess how well the proposed ECTT-PCSANN-WWPA approach, one statistic is the F1-score. It is computed in equation (32).

$$\text{F1score} = \frac{TP}{TP + \frac{1}{2}[FP + FN]} \quad (32)$$

3) Precision

Precision (P) is a metric which quantifies the count of correct positive prediction made. This is computed via following equation (33).

$$P = \frac{TP}{TP + FP} \tag{33}$$

4) Recall

Recall (R) is a statistic that measures how many correct positive forecasts there are out of all positive forecasts. The following equation (34), is used to measure it.

$$R = \frac{TP}{TP + FN} \tag{34}$$

B. Performance Analysis

Fig 2 to 5 shows the simulation outcomes of ECTT-PCSANN-WWPA. Next, the results are contrasted with the current ECTT-RNN, ECTT-DNN, and ECTT-MLPNN techniques..

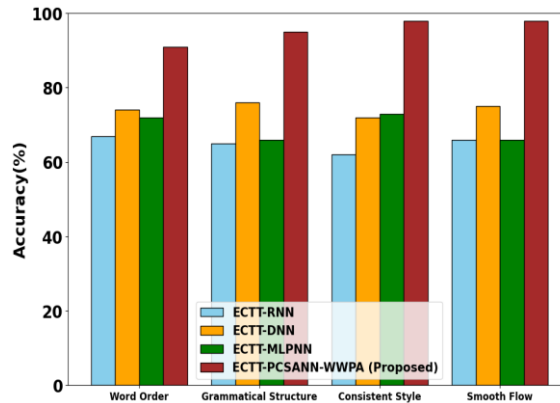


Fig 2: Comparison of accuracy between proposed and existing methods

Figure 2 compares the accuracy values of the suggested and current approaches. The effectiveness of the proposed method results in accuracy that are 50.54%, 20.77%, 35.96% higher for the classification of word order, 20.46%, 35.56%, 23.54% higher for the classification of grammatical structure, 21.44%, 30.73%, 18.41% higher for the classification of consistent style, and 21.46%, 30.86%, 15.45% higher for the classification of smooth flow when evaluated to the existing ECTT-RNN, ECTT-DNN, and ECTT-MLPNN methods correspondingly.

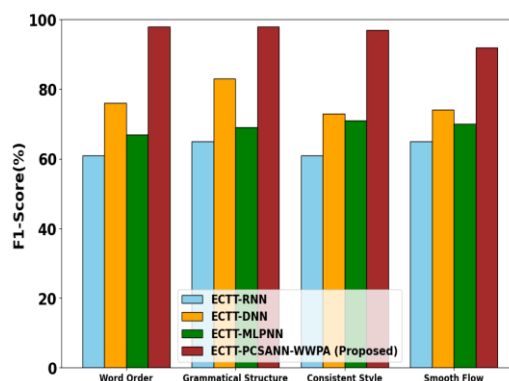


Fig 3: Comparison of F1-Score between proposed and existing methods

Figure 3 compares the F1-score values of the proposed and current techniques. The performance of the proposed technique results in f1-score that are 22.56%, 21.78%, 33.94%, higher for the classification of word order, 21.47%, 33.58%, 23.56% higher for the classification of grammatical structure, 21.45%, 30.77%, 18.42% higher for the classification of consistent style, and 20.45%, 30.88%, 15.46% higher for the classification of smooth flow when evaluated to the existing ECTT-RNN, ECTT-DNN, and ECTT-MLPNN methods correspondingly.

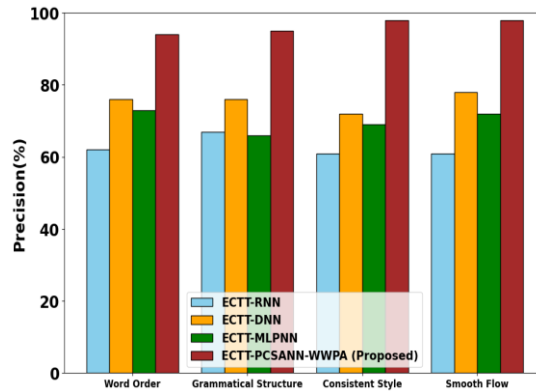


Fig 4: Comparison of precision between proposed and existing methods

Figure 4 compares the precision values of the proposed and current approaches. Here, a direct comparison with existing methods is offered to show how the proposed method's precision is higher. The proposed method provides for a more extensive analysis and has higher precision than existing methods due to its wider consideration of factors. The effectiveness of the suggested method results in precision that are 30.56%, 21.75%, 35.91% higher for the classification of word order, 21.48%, 33.57%, 23.53% higher for the classification of grammatical structure, 21.46%, 30.73%, 18.48% higher for the classification of consistent style, and 20.44%, 30.86%, 15.47% higher for the classification of smooth flow when evaluated to the existing RNN, DNN, and MLPNN methods correspondingly.

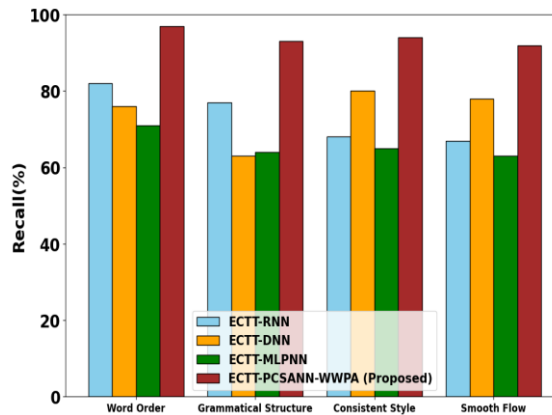


Fig 5: Comparison of recall value between proposed and existing methods

Figure 5 displays the recall value comparison between the proposed and current systems. The effectiveness of the proposed method results in recall that are 23.52%, 22.78%, 31.95% higher for the classification of word order, 22.46%, 31.57%, 22.57% higher for the classification of grammatical structure, 22.48%, 29.72%, 17.43% higher for the classification of consistent style, and 19.49%, 29.82%, 14.45% higher for the classification of smooth flow when evaluated to the existing ECTT-RNN, ECTT-DNN, and ECTT-MLPNN methods respectively.

V. CONCLUSION

In conclusion, this research harnesses the power of Pyramidal Convolution Shuffle Attention Neurological Network to significantly train the English-Chinese translation model. An essential first step is data collection where the data is acquired from UM-corpus news dataset which consists of large English-Chinese parallel corpus. The SEGBOT: Neural Text Segmentation method is used to segment the words from the dataset. The Iterated Square-Root Cubature Kalman Filter (ISCKF) is used to process the corpus dataset during pre-processing. The pre-processed data is fed to the classification where the corpus data is classified using Pyramidal Convolution Shuffle Attention Neural Network (PCSANN) into word order, grammatical structure, consistent style, and smooth flow. The English text is given as the input to the proposed algorithm to translate it into Chinese text using an encoder-decoder model based on Google's Transformer. The suggested approach is evaluated on the Python working platform and compared with existing methods. The proposed approach is examined in a variety of scenarios, including recall, accuracy, f1-score, and precision. The PCSANN classifier is utilized at the end of the system and shows that it is capable of precisely detecting accuracy, f1-score,

precision, and recall. Its accuracy has grown to 98% due to the precise identification of the recommended best region-expansion approach..

Acknowledgements:

This work was supported by Research on the Allocation of Cognitive Resources in Chinese and English Translation of College Students from 2022 Jiangsu Philosophy and Social Science Research Project of Higher Learning (No. 2022SJYB2401).

REFERENCE

- [1] Pei, Z. Z., & Chi, F. W. (1987). The two faces of English in China: Englishization of Chinese and nativization of English. *World Englishes*, 6(2), 111-125.
- [2] Tian, L., Wong, D. F., Chao, L. S., Quaresma, P., Oliveira, F., & Yi, L. (2014, May). UM-Corpus: A Large English-Chinese Parallel Corpus for Statistical Machine Translation. In *LREC* (pp. 1837-1842).
- [3] Chowdhary, K., & Chowdhary, K. R. (2020). Natural language processing. *Fundamentals of artificial intelligence*, 603-649.
- [4] Wang, H. (2019). The development of translation technology in the era of big data. *Restructuring Translation Education: Implications from China for the Rest of the World*, 13-26.
- [5] Stahlberg, F. (2020). Neural machine translation: A review. *Journal of Artificial Intelligence Research*, 69, 343-418.
- [6] Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T., & Maida, A. (2019). Deep learning in spiking neural networks. *Neural networks*, 111, 47-63.
- [7] Araabi, A., & Monz, C. (2020). Optimizing transformer for low-resource neural machine translation. *arXiv preprint arXiv:2011.02266*.
- [8] Baziotis, C., Haddow, B., & Birch, A. (2020). Language model prior for low-resource neural machine translation. *arXiv preprint arXiv:2004.14928*.
- [9] Al-Marouf, R. S., Salloum, S. A., AlHamadand, A. Q. M., & Shaalan, K. (2020). Understanding an Extension Technology Acceptance Model of Google Translation: A Multi-Cultural Study in United Arab Emirates. *International Journal of Interactive Mobile Technologies*, (3).
- [10] Md Abdur, R., Islamb, M. S., Hossainc, S., & Jiang, J. (2019). Exploring and Learning English: An Analysis of Baidu and Google Translation. *International Journal of Linguistics, Literature and Translation (IJLLT)*.
- [11] Slawski, M., Ben-David, E., & Li, P. (2020). Two-stage approach to multivariate linear regression with sparsely mismatched data. *The Journal of Machine Learning Research*, 21(1), 8422-8463.
- [12] Schwenk, H., Wenzek, G., Edunov, S., Grave, E., & Joulin, A. (2019). Ccmatrix: Mining billions of high-quality parallel sentences on the web. *arXiv preprint arXiv:1911.04944*.
- [13] Benkova, L., Munkova, D., Benko, L., & Munk, M. (2021). Evaluation of English–Slovak neural and statistical machine translation. *Applied Sciences*, 11(7), 2948.
- [14] Brynjolfsson, E., Hui, X., & Liu, M. (2019). Does machine translation affect international trade? Evidence from a large digital platform. *Management Science*, 65(12), 5449-5460.
- [15] Durrani, N., Dalvi, F., Sajjad, H., Belinkov, Y., & Nakov, P. (2019, June). One size does not fit all: Comparing NMT representations of different granularities. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 1504-1516).
- [16] Krüger, R. (2020). Explication in neural machine translation. *Across Languages and Cultures*, 21(2), 195-216.
- [17] Sun, M., Zhao, X., & Khan, A. (2022, April). Building a Chinese-English Parallel Corpus for Machine Translation: A Big Data-Driven Approach. In *International Conference on Multi-modal Information Analytics* (pp. 851-858). Cham: Springer International Publishing.
- [18] Yang, C. C., & Luk, J. (2003). Automatic generation of English/Chinese thesaurus based on a parallel corpus in laws. *Journal of the American Society for Information Science and Technology*, 54(7), 671-682.
- [19] Manoharan, D. J. S. (2021). Capsule network algorithm for performance optimization of text classification. *Journal of Soft Computing Paradigm*, 3(1), 1-9.
- [20] Akhter, M. P., Jiangbin, Z., Naqvi, I. R., Abdelmajeed, M., Mehmood, A., & Sadiq, M. T. (2020). Document-level text classification using single-layer multisize filters convolutional neural network. *IEEE Access*, 8, 42689-42707.
- [21] Xu, B. (2022). English-Chinese Machine Translation Based on Transfer Learning and Chinese-English Corpus. *Computational Intelligence and Neuroscience*, 2022.
- [22] Sun, Y. (2022). Analysis of Chinese machine translation training based on deep learning technology. *Computational Intelligence and Neuroscience*, 2022.
- [23] Xue, S., & Wang, X. An informative teaching model of English-Chinese translation in universities based on the perceptual machine model. *Applied Mathematics and Nonlinear Sciences*.
- [24] Zhao, X., & Jin, X. (2022). A comparative study of text genres in english-chinese translation effects based on deep learning LSTM. *Computational and Mathematical Methods in Medicine*, 2022.

- [25] Li, C. (2022). A study on Chinese-English machine translation based on transfer learning and neural networks. *Wireless Communications and Mobile Computing*, 2022.
- [26] Ruan, Y. (2022). Design of intelligent recognition English translation model based on deep learning. *Journal of Mathematics*, 2022, 1-10.
- [27] Zhao, L., Gao, W., & Fang, J. (2021). High-Performance English–Chinese Machine Translation Based on GPU-Enabled Deep Neural Networks with Domain Corpus. *Applied Sciences*, 11(22), 10915.
- [28] Gao, B. (2021). Analysis of the needs of English-Chinese translation courses and research on teaching strategies under the background of wireless communication and big data. *Wireless Communications and Mobile Computing*, 2021, 1-5.
- [29] Li, J., Chiu, B., Shang, S., & Shao, L. (2020). Neural text segmentation and its application to sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, 34(2), 828-842.
- [30] Johnson, T., Banu, S., & Moger, T. (2024). Forecasting-Aided State Estimation in Power Systems During Normal Load Variations Using Iterated Square-Root Cubature Kalman Filter. *Journal of Electronics and Electrical Engineering*, 35-49.
- [31] Chen, K., Wang, X., & Zhang, S. (2022). Thorax disease classification based on pyramidal convolution shuffle attention neural network. *IEEE Access*, 10, 85571-85581.
- [32] Alhussan, A. A., Abdelhamid, A. A., El-Kenawy, E. S. M., Ibrahim, A., Eid, M. M., Khafaga, D. S., & Ahmed, A. E. (2023). A Binary Waterwheel Plant Optimization Algorithm for Feature Selection. *IEEE Access*.