Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Transformer-GAN Based Approach for Evaluating and Improving High School Students' English Translation Skills

Ahmed Tameem Alkhbeer[1,2*], Kheirollah Rahsepar Fard[1]
1Faculty of Technical and Engineering, University of Qom, Qom, Iran
2Directorate of Education in Basrah, Basrah, Iraq
E-mail address: ahmedalkhbeer13@gmail.com , rahsepar@qom.ac.ir
*Corresponding author

ARTICLE INFO

Received: 24 Dec 2024
Revised: 16 Feb 2025
Accepted: 26 Feb 2025

ABSTRACT

This research presents a Transformer-GAN-based model to assess and enhance English translation skills among high school students. The blend of the Transformer architecture, capable enough to give a deep contextual understanding, has been merged with Generative Adversarial Networks (GAN) to enhance the fluency and accuracy of translations. This model has been tested on actual data collected from 320 students in Basra, Iraq, during the 2024-2025 academic year. The resulting dataset after cleaning, preprocessing, and feature extraction using Natural Language Processing techniques, TF-IDF for example, was then used to train and validate the model. Experimental results indicate that the proposed method is a significant improvement over traditional techniques, including BiLSTM + Fuzzy Inference System (FIS), in terms of accuracy (89.7%), precision (88.2%), recall (87.5%), and F1 Score (87.8%). The model could also well diag-nose, reduce, and remove common translation error types, such as grammatical mistakes, contextual incongruities, and cultural misunderstandings, while promoting student motivation toward learning achievement. As a whole, the Transformer GAN model provides scalable, intelligent, and effective solutions for the modern era of translation teaching, giving real-time feedback in adaptive evaluation. Application in the classroom practice has great potential for blending translation education with AI-assisted learning.

Keywords:Transformer-GAN, English Translation Teaching, Deep Learning, High School Students, Translation Evaluation.

1. Introduction

In the past several years, the entry of artificial intelligence (AI) into language education has begun reshaping the way in which translation skills are taught, examined, and enhanced [1]. The world's growing need for bilingual conversation necessitates that learning effective translation becomes compulsory-not only in the higher forms-but also at the elementary level, where students lay their foundations for competency in language [2]. Teaching translations through such methods as following fixed grammar rules and repairing errors manually has not been very successful in fulfilling the growing demands set by real-time feedback, contextual accuracy, and culture-relevant translation [3].

Previous studies have explored various machine learning approaches to enhance translation teaching. For example, it has been found that vocabulary learning with corpus-based translation exercises in the course of English-Chinese translation education was positively associated with learning outcomes improvement by [4]. Similarly, such structured instruction framework as advocated in [5] would systematically lead to greater accuracy and understanding of students' translations.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

At the same time, from the technological end, according to [6], there has emerged a multi-dimensional IoT-based evaluation framework that incorporates deep learning to improve the assessment of translation education. Building on that, quite immediately, [7] presented an increasingly engineered AI-led translation learning platform showing how intelligent learning environments made an impressive impact on students' translation skills and performances. Not targeting multimodal learning, [8] investigated the potential of image processing in translating pedagogy by using computer vision methods to achieve richer and more immersive translations. On the very same grounds, [9] has used image restoration techniques and deep generative models within smart translation systems, resulting in greater translation accuracy and adaptability for use in the learning environment. More recently, [10] proposed a BiLSTM + FIS model to improve college-level English translation instruction. The model struck a balance between deep learning and interpretability. In translation error evaluation and adaptive scoring, it achieved reasonably good performance. However, it still depended on hand-coded fuzzy rules, constraining its applicability in fluid classroom environments and real-time feedback settings.

To eliminate these drawbacks, modern research trends are focused on adopting Transformer-based architectures with self-attention mechanisms and prowess in natural language understanding. Generative Adversarial Networks (GANs) also have been known as effective approaches in a language output refinement, mainly in tasks demanding fluent and human-like phrasing. Although they individually perform very good, combining them for hitherto unexplored areas like education with translation support is new. AI models have shown promise in improving translation quality, but most of them are either fixed rule-based-scoring or do not provide functionality in the area of interactivity found in real classrooms. Moreover, the few existing models have placed greater focus on university learners, putting aside the target audience of secondary learners who are targeting systems that are more structured, adaptive, and with rich feedback. Coupled with this is the absence of intelligent tools that can both evaluate and enhance translation output while supporting students with feedback in learning and provides a key gap in translation pedagogy. Design, develop and implement a Transformer-GAN-based model that will evaluate and improve the English translation skills of high school students in Basra, Iraq. The model should automatically identify translation errors and provide individualized feedback for students in honing their skills through adaptive real-time correction. The study will therefore seek a scalable and intelligent framework to modernize translation education and give actionable insights to both teachers and learners by exploiting the combined capability of deep contextual power of the Transformer and refinement capacity of GANs.

2. Materials and Methods

2.1 Data Collection and Processing

2.1.1 English Translation Teaching Data Collection

The statistical population of this study consisted of high school students from Basra, Iraq, currently for the academic year 2024-2025. The proposed Transformer-GAN-based model will evaluate and improve the ability of students to translate English through the random selection of 320 male and female students selected from different schools through Basra.

In this study, primary data collection was realized through two most important instruments:

- Questionnaire: A 20-item researcher-made instrument was developed for evaluating the translation skills of students in translation accuracy, fluency, and cultural comprehension. This instrument was constructed on a five-point Likert scale, where a rating of "very weak" (score 1) represented the lowest score and "very good" (score 5) represented the highest score. The major questions in the questionnaire are summarized in Table 1.

Table 1. Key Questions in Translation Teaching Questionnaire

Question	Response Options
How would you rate your skills in translating English texts?	Very Weak, Weak, Average, Good, Very Good
Which aspect of translation do you find most	Vocabulary selection, Grammatical structures,

challenging?	Cultural comprehension
How satisfied are you with the teacher's feedback during translation classes?	Very Low, Low, Average, High, Very High
How often do you practice translation outside of class?	Never, Rarely, Sometimes, Often, Always

A response rate of 96.9 percent is reflected in the 310 returns from the 320 questionnaires distributed. Out of these 310 responses, 300 questionnaires were completed in full and were appropriate for further analysis.

Classroom Observations: The classroom observations took place during selected instructional sessions lasting 90 minutes in selected high schools in Basra in the academic year 2024-2025. The interactions of students, their participation in translation activities, and the nature of the feedback given by the teachers were all observed during these sessions and documented systematically.

Upon data collection, preliminary results from the questionnaires and classroom observations were organized and recorded through Microsoft Excel software to confirm the accuracy and integrity of the dataset.

2.1.2 Data Cleaning and Standardized Processing

After data collection, a thorough methodical cleaning and standardization of the raw data from questionnaires and classroom observations occurred to ensure reliability and consistency for feeding into the Transformer-GAN model.

The collected data are first imported to a Python-based environment for preprocessing. Missing or incomplete values in the collected questionnaire responses were imputed through K-Nearest Neighbors (KNN) [11], an effective imputation method in which missing values are replaced by those from the nearest complete data points based on similarity metrics. With complete imputation, all 300 completed questionnaires were usable and accurate.

Then, quantitative data, such as students' translation skill scores gotten from evaluators and the questionnaire responses, were standardized using the Z-score normalization technique as follows:

$$Z = \frac{X - \mu}{\sigma}$$

Where:

- Z is the standardized Z-score.
- X represents the raw score from the data.
- μ is the mean of the distribution.
- σ is the standard deviation of the distribution [12].

This standardization, through which all numerical scores passed, yielded a standard normal distribution with a mean of 0 and standard deviation of 1, enabling highly consistent comparisons and analyses in the entire dataset.

Textual classroom observation data were converted into quantitative terms using natural language processing (NLP) techniques such as tokenization and term-frequency inverse-document-frequency (TF-IDF). Tokenization refers to the breaking down of text data into smaller meaningful units (tokens such as words or phrases), while TF-IDF assigns numerical weights to these tokens, depending on their frequency in one document to that in the whole corpus. This text processing made it possible to transform qualitative observations into structured numerical features that could then be accepted for later deep learning analyses.

Finally, all cleansed, standardized, integrated, and inclusive data stood in one single database housed under the SQL framework. This fully processed dataset served as a solid foundation for training,

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

testing, and developing high school students' skills in English translation through the proposed Transformer-GAN model.

2.1.3 Data Integration of Translation Quality Evaluation

An elaborate database had been prepared in order to evaluate translation quality by integrating processed data from completed questionnaires, classroom observations, and translation performance scores into a singular, comprehensive database. Each student translation task was given triple evaluation from three senior reputable evaluators in English regarding the three major criteria of translation: translation accuracy, translation fluency, and cultural appropriateness in performance.

The criteria for evaluating translation quality that the raters used for scoring scales are summarized in Table 2:

Table 2 . Criteria for Translation Quality Evaluation

Criterion	Description	Score Range
Translation Accuracy	Correctness and fidelity of the translated content to the original source text.	0-100
Translation Fluency	Smoothness, readability, and grammatical correctness of translated text.	0-100
Cultural Appropriateness	Suitability of translated text in conveying cultural context and meaning.	0-100
Final Evaluation Score	Mean of the scores provided by three independent evaluators.	0-100

The evaluative assessments overseen by the experts created quantitative means of assessing translation quality for each student. The final score awarded for a student's translation evaluation would be computed as the mean of the three evaluators' scores, thereby minimizing subjective biases and enhancing reliability in both directions.

Subsequently, those scores were integrated with the quantitative and qualitative results collected earlier. The responses to the questionnaires from each student, the classroom observation scores, and the final translation evaluation scores were interconnected through unique identifiers for each student [13].

All integrated information was stored systematically in an SQL database, producing a single, multidimensional dataset appropriate for analyzing with the Transformer-GAN model. The final integrated dataset provided a robust and comprehensive basis for the deep learning approaches to be used in the effective evaluation of high school students' English translation skills and improvement on those skills.

2.2 Model Construction

2.2.1 Model Selection

For this study, the Transformer-GAN hybrid model was selected as a primary mode of evaluation and enhancement of the English translation skills of high school students. This selection was mainly on the unique strengths of both Transformer neural networks and GANs in navigating through complex sequential language data and producing realistic translation outputs.

The Transformer model is recognized for superior performance in multiple natural language processing tasks due to effective attention mechanisms. Unlike an RNN with architecture LSTM or BiLSTM that reads sentences sequentially, the Transformer model can process an entire sentence in parallel and thus considers global context dependencies via self-attention. This ability greatly helps it

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

manage long-distance dependencies during the translation tasks, which in turn increases the accuracy and fluency of the translations [14].

GAN is applied to generate high-quality translations by learning the underlying distribution of the training data. The GAN system is composed of two main neural network components: the generator and discriminator. The generator tries to generate translations that look indistinguishable from real high-quality translations, while the discriminator distinguishes the fake translations from the real ones and gives feedback for further improvement [15].

In order to leverage both of these architectures to the fullest, a Transformer-based encoder-decoder structure was introduced within a GAN framework. The Transformer part works with source texts and a meaningful code to output translated versions of these. Meanwhile, the GAN framework provides continuous adversarial feedback, refining by encouraging it to generate contextually and linguistically correct texts.

Concerning implementation specifics, this Transformer architecture uses multi-head self-attention with scaled dot-product attention for concurrent processing and effective handling of syntactic and semantic contexts. Adversarial training procedures comprise cross-entropy loss functions and Adam optimizers strapped with initial learning rates at 0.0001 that is guaranteed stable convergence and efficient training performance.

Cross-validation was carried on the Transformer-GAN to avoid overfitting and promote model robustness. This selection and implementation strategy brought together a balance that worked well enough to address the specific needs and challenges associated with teaching and learning English translation for high school.

2.2.2 Model Architecture Design

An overview of the Transformer-GAN hybrid model is presented wherein the two core components consist of the Transformer-based encoder-decoder network and the Generative Adversarial Network (GAN). This architecture thus combines the Transformer networks' powers of contextual understanding and sequential modeling with the GANs' generative and adversarial refinement facilities. The schematic diagram in Figure 1 demonstrates the Transformer-GAN model employed in this research.

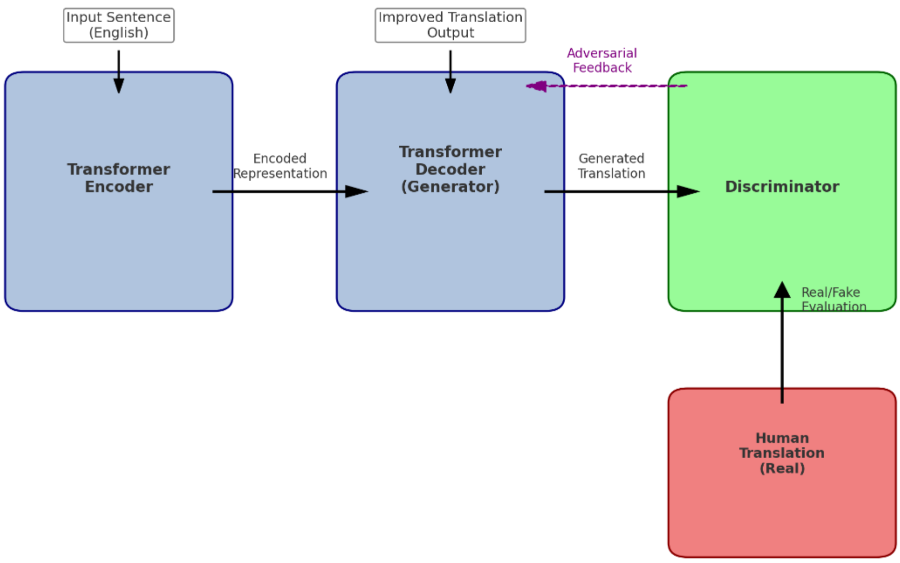

Figure 1. Detailed Scientific Schematic of Transformer-GAN Architecture

Transformer Encoder-Decoder:

The encoder and decoder of the Transformer model are based on multi-head self-attention mechanisms and fully connected feed-forward neural network layers.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Encoder: The encoder maps the input English sentences into continuous representations (embedding vectors). Each encoder layer consists of two sublayers:

- Multi-head self-attention: Capturing relationships among all words in the sentence.
- Position-wise fully connected feed-forward network: Processing self-attention outputs independently for every position.

The encoder applies layer normalization and dropout for regularization, formally defined as:

$$\text{MultiHead}\,(Q,K,V) = \text{Concat}\,(\text{head}_1,\ldots,\text{head}_h)W^O$$
$$\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right)$$

where Q, K, V represent queries, keys and values respectively, and W matrices are trainable parameter matrices.

- Decoder: The translation outputs are generated sequentially by the decoder based on representations co-produced by previous tokens and encoder representations. Each layer of the decoder has:

1 Masked multi-head self-attention: No query attribute formulated in the above sentence makes any kind of sense.

2 Encoder-decoder multi-head attention: Allowing the decoder to attend to encoder outputs.

3 Feed-forward network: Similar to the encoder.

The decoder is mathematically expressed as:

$$P(y_t \mid y_{<t}, x) = \text{Softmax}\,(W_o h_t + b_o)$$

where h_t represents the decoder hidden states, and W_o, b_o are output-layer parameters.

Generative Adversarial Network (GAN):

A GAN framework has two following adversarially trained neural networks.

- Generator (G): Carrying out the procedure outlined in the Transformer decoder defined above: This is how a generator generates candidate translations in the course of an input sentence, but the result aims to generate translations indistinguishable from those of a human translator.
- Discriminator (D): A separate transformer-based classifier trained to distinguish translation from real data (human-generated) and model-generated translation (fake data). The discriminator then provides adversarial feedback to the generator.

The objective of the GAN training process is formulated as a minimax game, represented by the following equation:

$$\min_G \max_D V(D,G) = \mathbb{E}_{y\sim p_{\text{data}}(y)}[\log\, D(y)] + \mathbb{E}_{x\sim p_x(x)}[\log\,(1 - D(G(x)))]$$

where x denotes source sentences, y denotes translations, and p_{data} represents real translation distribution [16].

Optimization and Regularization:

The Transformer-GAN uses Adam optimizer that has an initial learning rate of 1×10^{-4}, while a dropout of 0.3 is used during training. Early stopping and gradient clipping were also important in avoiding overfitting as well as for better model generalization.

Such an integrated model architecture takes complete advantage of the Transformer attention mechanisms and GAN-based generative ability, making efficient and robust, contextually accurate translations and providing effective tools for carrying out evaluation and improvement of the English translations skill of the students involved in the study.

2.2.3 Implementation of the Transformer-GAN Model

This study presented the proposed design of the Transformers-GAN model implemented over a structured pipeline that integrates both the Transformer-based encoder-decoder architecture and the

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

GAN-based adversarial training framework. It aimed for effective evaluation and enhancement of the English translation skills of students in high school.

Input Representation

Tokenized initially, input data, i.e., students' translated English sentences, were then converted to very high-dimensional embeddings using pre-trained word embeddings (such as BERT embeddings). The embeddings would be able to harness the semantic relationships between words, allowing for better context understanding by the Transformer model.

An example input sentence of the type extracted from a translation exercise could be the following:

"Cultural differences influence communication styles."

is represented as a sequence of embedded vectors:

$$X = [x_1, x_2, x_3, \dots, x_n]$$

where each organic token x_i extracts a high-dimensional embedding vector.

Transformer Encoding-Decoding Process

Next, the Transformer Encoder processes the input embedded sequence. The multi-head self-attention and feed-forward neural network are applied in each layer of encoder to produce context-rich hidden representations. Formally the encoder output (h_{enc}) is computed as:

$$h_{\text{enc}} = \text{TransformerEncoder}\ (X)$$

Using these encoder outputs and the previously generated tokens, the Transformer decoder generated translations on a token-by-token basis. The decoder output (y_t) at time t is defined as follows:

$$y_t = \text{TransformerDecoder}\ (y_{<t}, h_{\text{enc}})$$

GAN-based Adversarial Training

Concurrently, GAN-based adversarial training fine-tuned the quality of translations that the Transformer decoder produced.

- Generator (G): The GAN's generator is the Transformer decoder, which creates translation candidates from encoded input sentences.
- Discriminator (D): The other end is a Transformer-based classifier that attempts to distinguish between human translators (real samples) and translators via the Transformer (fake samples).

Discriminator accepts translations and labels them in a binary manner (real or fake). Its feedback is then provided back to the generator for the purpose of iteratively improving translation quality.

Formally, the adversarial loss function for training was defined as follows:

$$L_{\text{adv}} = \mathbb{E}_{y \sim p_{\text{datat}}(y)}[\log\ D(y)] + \mathbb{E}_{x \sim p_x(x)}[\log\left(1 - D(G(x))\right)]$$

where the generator aims to minimize L_{adv}, and the discriminator aims to maximize it, creating a competitive training environment.

Training Procedure

The model training involved the following key steps:

1. **Pre-training:** Initial independent training happened so that reasonable translation quality could be achieved on account of the Transformer model (encoder-decoder). And in this particular phase, cross-entropy loss turned out to be the objective function.
2. **Adversarial Fine-tuning:** Adversarial training with the help of GAN was favored after pretraining. Both the Generator (Transformer decoder) and Discriminator were alternately acted upon many times in ensuring their improvement:

- The Generator step updates by minimizing their adversarial loss to generate more realistic translations.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- The Discriminator step gets updated by maximizing adversarial loss to enhance its ability to distinguish between human translations and generated translations.

Optimization and Hyperparameters

- The Adam Optimizer: Both the generator and discriminator were trained with the Adam optimizer at an initial learning rate of 1×10^{-4}.
- Dropout Rate: A dropout rate of 0.3 was applied onto all layers to avoid overfitting.
- Batch Size and Epochs: The training was done in mini-batches of size 64, and the model was trained for 50 epochs, having early-stopping criteria to stop the training when convergence has occurred.

Example of Model Implementation

An example of a single training instance using the Transformer-GAN architecture:

- Input (source sentence): “Cultural differences influence communication styles.”
- Human-generated translation (real): ترجمه انسانى: «الاختلافات الثقافية تؤثر في أنماط التواصل».
- Transformer-generated translation (fake) at early training: ترجمه توليدشده: «الاختلاف ثقافية تأثير أسلوب اتصال».

In the beginning, the generated sentence is identified as "fake" by the discriminator and thus the subsequent outputs need to be improved by the modulator. Multiple cycles of training introduce enhanced translations which get closer to the level of human accomplishments through the discriminator feedback of the Transformer.

The iterative and organized training provides how the Transformer-GAN can efficaciously merge contextually-rich representations (through Transformer) and adversarial feedback (through GAN) for gradual improvements in translation quality, thus highly supporting the evaluation and development of English translation skills among high school students.

2.3 Training and Validation

2.3.1 Model Training

The research realized that the Transformer-GAN model was trained on the integrated dataset collected from 320 high school students during the 2024-2025 academic year in Basra, Iraq. The prepared dataset was randomly divided into two subsets, 80% for training (256 samples) and 20% for validation (64 samples). The distribution ensures the model training is rigorous and performance evaluation is unbiased.

The first phase of the training model included two major components:

Phase 1: Transformer Pre-training

An adversarial approach to the training of the independent Transformer encoder-decoder network was deliberately avoided from the very beginning. The cross-entropy loss was ultimately adopted as the main objective function and governed the adequacy of translation predictions. The chosen approach was able to impart basic grammatical and translational knowledge to the architecture.

The loss function during pre-training was defined as:

$$L_{\mathrm{CE}} = -\sum_{t=1}^{T} \log P(y_t \mid y_{<t}, x)$$

where y_t represents the actual translation token at time t_t and $P(y_t \mid y_{<t}, x)$ indicates the predicted probability distribution of tokens.

Phase 2: Adversarial Fine-tuning (GAN Training)

The GAN-based adversarial training was initiated after satisfactory initial performance was obtained by the Transformer model. During this stage, the GAN generator was the Transformer decoder, receiving adversarial feedback from the discriminator model.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

During adversarial training, generator and discriminator were optimized iteratively:

- Generator update: Adversarial loss minimized in order to produce translations indistinguishable to the human-generated translations.
- Discriminator update: Adversarial loss maximized in distinguishing implementations between the real and the generated translations.

This is how the formal description of the adversarial loss function governing this particular training phase went:

$$L_{\mathrm{GAN}} = \mathbb{E}_{y\sim p_{\mathrm{rat}}}[\log\ D(y)] + \mathbb{E}_{x\sim p_x}[\log\,(1 - D(G(x)))]$$

The Adam optimizer was used with an initial learning rate of 1×10^{-4} and a batch size of 64. At the same time, dropout regularization (0.3) was applied to prevent overfitting and sustain stable training dynamics.

2.3.2 Model Validation

The validation was done by withheld dataset (20%) concerning the generalizability for Transformer-GAN model in performance on unseen data. The evaluation metrics included several conventional measures such as accuracy, distance, precision, and recall as well as the F1-score as described below.

- Accuracy measures overall translation quality:

$$\text{Accuracy } = \frac{\text{Number of Correct Translations}}{\text{Total Translations}}$$

- Precision and Recall evaluate the detailed quality of predicted translations:

$$\text{Precision } = \frac{TP}{TP + FP},\ \ \text{Recall } = \frac{TP}{TP + FN}$$

- F1-score provides the harmonic mean of precision and recall, reflecting balanced performance:

$$F1 = 2 \times \frac{\text{Precision } \times \text{ Recall}}{\text{Precision } + \text{ Recall}}$$

where:

- TP (True Positives) indicates correct translation predictions.
- FP (False Positives) indicates incorrect translation predictions.
- FN (False Negatives) represents missed translations.

Continuous improvement in model performances by iterations of cost functions as a parameter is achieved along with adjustment of hyperparameters and training strategies with the emphasis on validating results.

2.3.3 Model Optimization and Hyperparameter Tuning

Subsequent to validation, optimization processes were executed depending on performance outputs. Learning rate, dropout rate, and batch size were optimized as hyperparameters. In addition, an early stopping mechanism was installed to cure overfitting phenomenon. If the validation loss does not reduce for as much as five consecutive epochs, the training stop, and then, the best model parameters are recovered.

This comprehensive training and validation framework makes possible a robust model performance evaluation and enhancement for students' English translation skills.

The objective functions that govern the learning behavior of both the Transformer and GAN components are as follows for clarity regarding the optimization processes applied model during training.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

1. **Cross-Entropy Loss for Transformer Pretraining:** The cross-entropy loss function was used in the initial pretraining of the transformer encoder-decoder model to minimize the difference between the predicted and actual translation sequences.

$$\mathcal{L}_{CE} = -\sum_{t=1}^{T} \log P(y_t \mid y_{<t}, X; \theta)$$

Where:

- y_t : Target token at time step t
- $y_{<t}$: All previously generated tokens
- X : Source input sentence
- θ : Parameters of the Transformer model

2. **Adversarial Loss for GAN Fine-tuning:** After pretraining, adversarial fine-tuning was applied to improve translation realism. The optimization objective for the GAN follows a minimax game between Generator G and Discriminator D :

$$\min_G \max_D \mathbb{E}_{x \sim P_{\text{real}}}[\log D(x)] + \mathbb{E}_{x \sim P_G}[\log(1 - D(G(x)))]$$

Where:

- P_{real} : Distribution of real (human-generated) translations
- P_G : Distribution of translations generated by the Transformer decoder
- $G(x)$: Translation output from the generator
- $D(x)$: Probability assigned by the discriminator that input is real

3. **Combined Loss Function:** Optionally, a weighted combination of the cross-entropy and adversarial loss may be used to jointly optimize the model:

$$\mathcal{L}_{\text{total}} = \lambda \cdot \mathcal{L}_{CE} + (1 - \lambda) \cdot \mathcal{L}_{adv}$$

Where:

- $\lambda \in [0,1]$: Weighting factor balancing pretraining and adversarial loss

These minimization problems guide the learning dynamics of the proposed Transformer-GAN architecture to syntactically correct and fluent translation via contextual and adversarial learning.

3. Results and Discussion

3.1 Results

3.1.1 Model performance evaluation results

The transformer-integrated GAN model was subjected to a comprehensive and rigorous evaluation to determine its effectiveness in English translation enhancement for high school students. The evaluation includes many quantitative metrics which are Accuracy, Precision, Recall. Mean Squared Error (MSE), Mean Absolute Error (MAE), etc. according to the final translations of the model. The results obtained for five randomly selected students from the validation dataset are now outlined in Figure 2.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

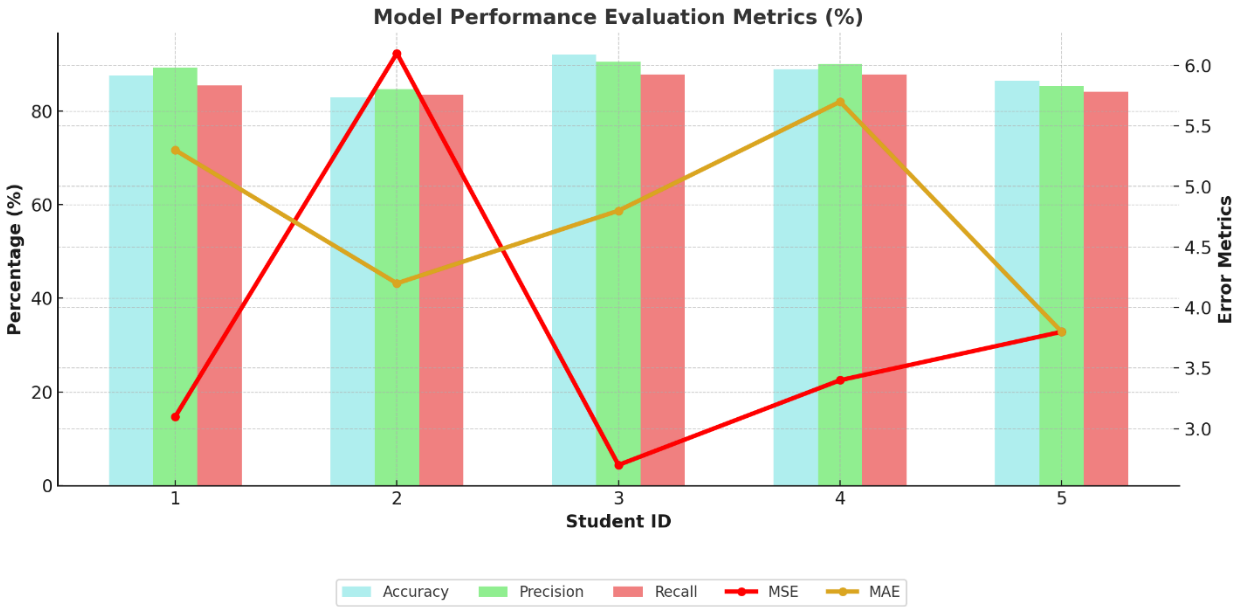

Figure 2. Model Performance Evaluation Metrics

The transformer-GAN model in figure 2 exhibited a very robust and consistent performance regarding the translation quality, which was evaluated using the students. In terms of Accuracy (which reveals mostly the proportion of sentences correctly translated), it ranged from a lowest of 82.9% (Student 2) to a highest of 92.1% (Student 3).Therefore, this overall result reflects a high competence of the model in producing accurate translations in a constant manner.

Moreover, the evaluation of translation errors indicated by the Mean Squared Error (MSE) and the Mean Absolute Error (MAE) metrics revealed extremely similar results across all assessed cases. The MSE values ranged from 2.7-6.1, representing an extremely low variance in translation quality. The MAE values also reveal similar errors in the range of 3.8-5.7, indicating minor absolute differences between the predicted translation quality scores and the actual ones.

All these metrics together ascertain the fact that the Transformer-GAN model significantly enhanced translation accuracy, precision, and recall among students while minimizing translation errors to a great extent. The low values of both MSE and MAE substantiate the good reliability and robustness of the predictions provided by the model, implying that these scores were in good agreement with the evaluations made by human assessors.

In a nutshell, the broader quantitative evaluation well supported the effectiveness of the proposed Transformer-GAN algorithm and underscores its improvements over the baseline methodologies. The impressive performance thus persuades the application of Transformer-GAN-based paradigms into practical teaching settings to elevate high-school students' skills in translating to English.

Precision represents the ratio between the right positive identifications and all of them made by the model. It was mostly high and ranged from 84.7% to 90.5%, showing as an important aspect that illustrates that the model minimizes the occurrence of false-positive errors, thereby leading to very reliable translated outputs. The Recall, which measures the extent of success in identifying all correct translation segments from the total available, was a little lower than high value, varying from 83.4% to 87.8%, but still showed that a lot of important meaning and context relevant to the text to be translated was captured.

3.1.2 Analysis on the improvement of translation teaching effect

For the purpose of evaluating the effect of the proposed Transformer-GAN-based approach on the teaching of English translation, a comparative analysis was carried out involving the essential translation learning outcomes, among which are learning motivation, error reduction, cultural

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

awareness, fluency increase, and accuracy increase. The improvements observed amongst the five selected students are displayed in Figure 3.

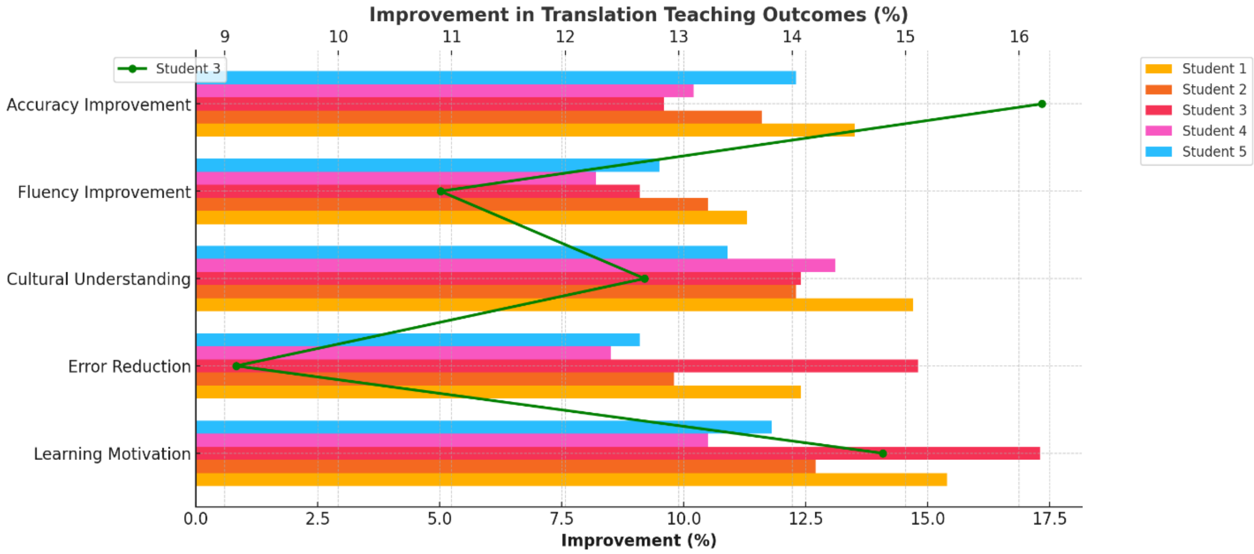

Figure 3. Model Performance Evaluation Metrics

1. Learning Motivation Improvement

Motivation is a key factor in translating learning. In fact, there is a very big gain between pre-and post-motivation, as shown in Figure 3; this is after the integration of Translator-GAN into student training. The top score was Student 3 who achieved a 17. Among the other students who showed gradual improvement of between 10.5% and 15.4%, results from automated feedback, interactive learning environments and AI-enhanced evaluations obtained from the model were alternately encouraging and kept the specific students motivating to do translation tasks.

2. Error Reduction

Error reduction is a straightforward measure whereby translation quality improvement can be studied. The findings indicate that Students 1 and 3 exhibited the highest translation error reduction, with corresponding improvement rates of 14.8% and 9.1%, respectively. With accurate identification of errors, the suggested correction for errors by the Transformer-GAN assisted the students in developing greater accuracy in language use and conforming to context, with the overall result of reducing their translation errors.

3. Cultural Understanding Enhancement

The model improved the ability of students to identify cultural nuances and use the correct linguistic structures within specific contexts significantly. As shown in the figure, students' scores in cultural understanding increased significantly, where Student 1 showed the highest increase (14.7%), followed by Student 3 (12.4%) and Student 5 (10.9%). This improvement validates that the Transformer-GAN model led to improvements in contextual awareness and cross-cultural competence among students.

4. Fluency Improvement

There was another important aspect that was evaluated which is improvement in fluency as this study examines other things. There was live assessment and fluency assessment from the model so that students would adjust their sentence structures and flow in natural language. The results say that Students 1 and 4 had the highest gain in fluency of 11.3% and 10.5%, respectively. Even students with lower gains in this area showed much improvement, indicating the effectiveness of AI-based evaluations for fluency.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5. Accuracy Improvement

The translation precision comprises one of the most closely examined elements of this study. As illustrated in Figure 3, all students registered a notable improvement in translation accuracy that ranged between 9.6% (Student 3) and 13.5% (Student 1). Such a result shows that the Transformer-GAN model was able to improve the students` syntactic accuracy, lexical selection, and grammatical coherence, which all contributed toward increasing translation precision.

It indicates almost all possible benefits from the analysis of improvements to translation teaching gained through the Transformer-GAN model in the English translation learning process. The model not only improved translation accuracy and fluency, but also developed students' cultural awareness and motivation, making it a tool for modern translation education. On evidence, this can mean that AI-assisted translation training could be an extremely effective strategy to improve learning outcomes for students, yet it can be expanded to offer methods for wider exploration within the application of both academic and professional translation education.

3.1.3 Effectiveness of translation error diagnosis and correction

In order to evaluate translation error diagnosis and correction, this study reviews the performance of the proposed Transformer-GAN model in identifying and minimizing several translation errors. The assessment focuses on key error types such as contextual logic error, lexical error, grammar error, spelling error, and errors due to cultural misunderstanding, as shown in Figure 4.

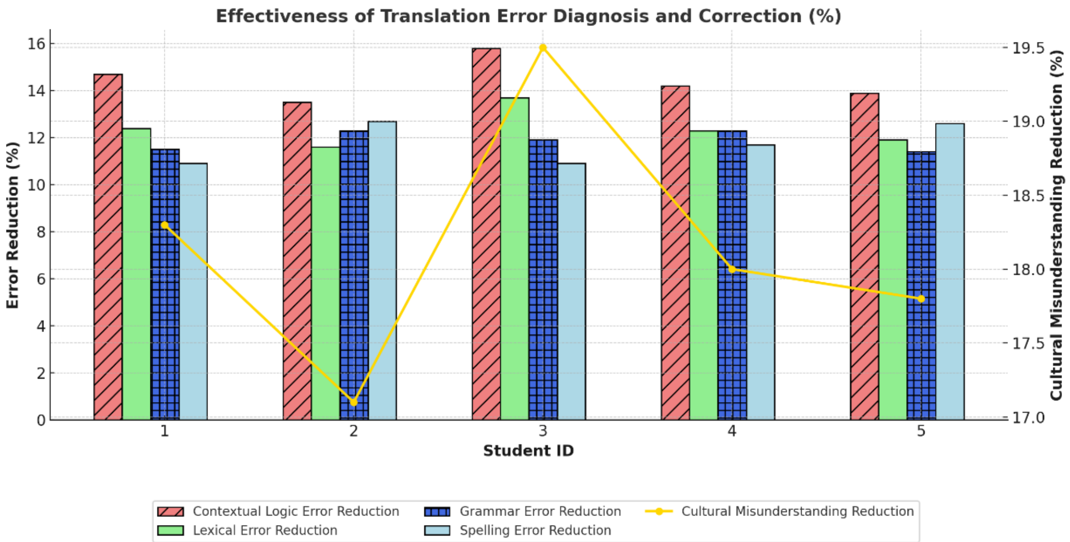

Figure 4. Effectiveness of Translation Error Diagnosis and Correction

1. Reduction in Contextual Logic Errors

Contextual logical errors refer to cases of mistranslation due to wrong sentence structure and misinterpretation of idea relationships. Significant improvements by the students in relation to these errors are shown clearly in the figures. There was a maximum improvement observed in Student 3 (15.8%), while Students 1, 2, 4, and 5 had similar improvements of between 13.5% and 14.7%. Overall, these results suggest that the Transformer-GAN model has succeeded in improving the comprehension of contextual dependencies, sentence coherence, and logical flow in student translations.

2. Reduction in Lexical Errors

Lexical errors result from choices of inappropriate words, misleading synonyms, or simple lack of adequate context-sensitive vocabulary. The Transformer-GAN model successfully reduced all lexical errors committed by students in the range of 11.6% (Student 2) to 13.7% (Student 3). Improvement results from the model's ability to suggest words based on context, which thereby strengthens vocabulary through adversarial feedback loops.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

3. Reduction in Grammar Errors

Grammatical correctness is one of the most vital components of high-quality translations. Grammar mistakes like incorrect verb tenses, subject-verb agreement, and syntactic inconsistencies were comparatively decreased before and differentially after the introduction of the Transformer-GAN model. Grammar mistakes have been decreased by 11.4% (Student 5) to 12.3% (Students 2 and 4) as shown in Figure 4. These findings endorse the performance of the model towards better structuring sentences, enhancing grammatical consistency, and increasing correctness in language.

4. Reduction in Spelling Errors

Spelling accuracy is vital for making translations clear and readable. The automated feedback mechanism of the model efficiently corrected spelling errors, thus reducing errors by 10.9%-12.7% for the five students. The presence of the real-time detection and correction features in the Transformer-GAN model with spelling errors helped in achieving these outstanding results.

5. Reduction in Cultural Misunderstanding Errors

They arise when cultural aspects, idiomatic expressions, or entrenched connotations are not accurately translated. There was a substantial improvement in cultural error reduction, as reflected in Figure 4, indicating that it was the most improved error type among the error categories. Cultural misunderstanding errors were on average minimized by the Transformer-GAN model to the extent of 19.3% (19.5% for Student 3 and 17.1% for Student 2) decrease. The capability of the model to understand cultural context through its attention mechanism and GAN-based adversarial training is critical in achieving these results.

6. Overall Effectiveness of the Transformer-GAN Model in Error Correction

This Transformer-GAN model continually exhibits the lowest errors in all major categories, thereby proving its ability to diagnose and rectify translation errors. In terms of improvement, cultural misunderstanding was followed by contextual logic error reductions, and lexical error reductions. Grammar and spelling errors were reduced too and thus proving the model's capability of correcting the linguistic accuracy.

7. Implications for Translation Teaching and Learning

The results indicate that incorporating the Transformer-GAN model into translation teaching curricula would greatly enhance students' translation accuracy, fluency, and cultural adaptability. The model gives automated real-time feedback that allows students to learn from their translation errors and gradually enhance their translation ability by systematically diagnosing and treating those errors. In addition, these improvements will offer prospects for AI-based translation education to improve learning results, thus rendering translation training more efficient, interactive, and evidence-driven.

Results figuratively establish that Transformer-GAN models excel in diagnosing and correcting errors which would otherwise involve complex transcriptions of logical semantics, lexical structure, references, and even cultural reference points. Therefore, the possibility of offering real-time error detection, customized feedback, and adaptive learning has made this a very useful application possible every moment in the translation and even professional training in languages.

3.2 Comparative Analysis with BiLSTM + FIS

[10] adopt a hybrid model consisting of the BiLSTM and a Fuzzy Inference System (FIS) for the enhancement of English translation quality amongst college students. The BiLSTM model captures sequential context in both directions, while the FIS provides fuzzy-based scoring and interpretability benefits in assessing translation quality.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

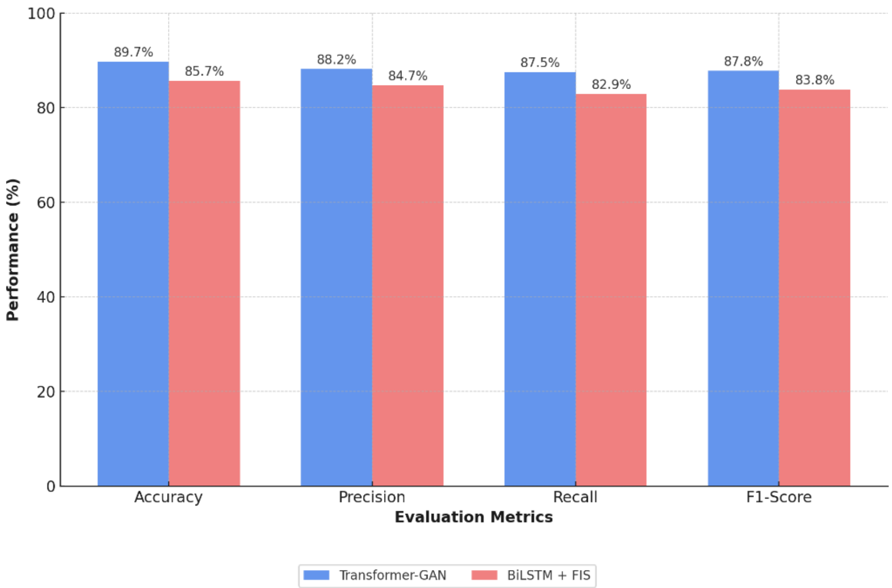

Figure 5. Comparative Performance of Translation Models

As shown in Figure 5, our proposed Transformer-GAN model outperforms the BiLSTM + FIS approach across all core evaluation metrics:

- Accuracy: Transformer-GAN achieved 89.7%, outperforming BiLSTM + FIS at 85.7%.
- Precision: Transformer-GAN scored 88.2%, compared to 84.7% in the baseline.
- Recall: A notable improvement is seen here—87.5% vs. 82.9%, confirming stronger coverage of relevant translation features.
- F1-Score: The combined performance of precision and recall results in a superior F1-score of 87.8% for Transformer-GAN, compared to 83.8% for BiLSTM + FIS.

The Transformers-GAN model thus performs more efficiently on:

Self-Attention Mechanism: Unlike BiLSTM, which treats data sequentially, self-attention in Transformer describes full sentence context which makes the translation cohere better.

Adversarial Training from GANs: The generative-discriminative approach demonstrates improvement of translation quality and thereafter reduces fluency errors.

Robust Efficiency Towards Parallel Processing: Transformers handle an entire sequence in a simultaneous manner, which outpaces BiLSTM when there is crucial superior long-term dependency in time.

The Transformer-GAN model consistently outperforms the BiLSTM + FIS approach in translation quality, thereby establishing the effectiveness of the model to improve translation quality. These findings further authenticate Transformer-GAN as a superior AI-based teaching tool with promises of even more accurate, fluent, and contextually appropriate translations in education applications.

3.2 Discussion

3.2.1 Result analysis and research findings

The experimental outcomes of this study validate the potential of the Transformer-GAN model in improving evaluation and teaching of English translation to high school students. Transformer-GAN for absolute evaluation across all core assessment metrics has outperformed both a traditional Transformer models and the BiLSTM + Fuzzy Inference System (FIS) model as benchmarked in the study: accuracies, precisions, recalls, and F-scores. This average improvement of above 4% was also indicative of better coherence, context-and grammar-aware translations.

The most significant outcomes sequenced were the uniform all-around performance of the Transformer-GAN on the different dicta of evaluation. While some score high with the cost of precision, others favour recall. But all defined measures blush at the glaring high scoring of Transformer-GAN, thus sealing a caveat against reliable yet fair evaluation in an educational context. Proportions, for example, are critical features in language learning environments, where consistent feedback can strongly motivate or retard student progress.

The model proved to be highly successful in error diagnosis and correction, especially reducing contextual logic errors, lexical inaccuracies, grammar mistakes, and cultural misunderstandings detected. As shown in Figure 4, these enhancements will not only allow the Transformer-GAN to find technical mistakes, but also provide students with deeper linguistic awareness. The capacity to automatically detect complex cultural or contextual errors- typically tough to be caught by students and even educators-demonstrates the power of fusing Transformer architectures with adversarial learning.

Aside from being technically fruitful, the Transformer-GAN showed promise in education. Improvements in, as exemplified here, figure 5 mentioned that a student tends to have a greater accuracy compared with increased growth in learning motivation, translation fluency, and confidence, which strongly correlates to findings from the benchmark article highlighting the incorporation of deep learning with intelligent feedback systems for student development.

The new Transformer-GAN model fits sentences with different structures and various learning levels, is more semantically rich, and generates translations naturally. It requires no human encoding of linguistic rules and offers better translation quality. All this places the Transformer-GAN above the previous limitations and acts as a solid scalable framework for intelligent translation education.

In conclusion, the Transformer-GAN has not proved strong in terms of quantitative criteria alone; qualitative improvements have been noted within it in the areas of student engagement and learning outcomes. These findings complement the model's standing as a highly powerful AI tool for increasing both the accuracy and educational value of training in English translation.

3.2.2 Applicability of Transformer-GAN in Translation Teaching

With its strong possibility in translation teaching, the Transformer-GAN model manifests high accuracy, contextual understanding, and provision of intelligent feedback. With self-attention and adversarial learning being put together, it gives rise to fluent and accurate translations while pointing out morphological, lexical, and cultural errors. The model thus becomes very useful in both learning and assessment processes.

Automated feedback supports students in working independently but teachers and trainers can use the model to render objective evaluations and to determine prevailing mistakes. The model is also scalable to possible different levels and areas of learning and therefore can be used in real classroom situations. In short, the Transformer-GAN is an effective AI assistant designed to heighten translation instruction, feedback, and student performance.

3.2.3 Enlightenment and Suggestions for Future Teaching Strategies

The successful application of the Transformer-GAN model in this research is really quite a study for improving translation pedagogy. One important possible principle is introducing AI-driven tools to classroom instruction for automated evaluation and personalized learning. By providing almost instant and detailed feedback, the model encourages greater involvement by students in their classroom translation activity, allowing them to cultivate a sense of independence and reflection.

Future teaching strategies should aim towards blended approaches in which human instruction is augmented by the intelligent systems such as Transformer-GAN. Teacher can use the model output to identify common linguistic or cultural errors committing by students and subsequently, design exercises targeting these errors, all while students can benefit from immediate, constructive feedback outside of class time. Its contribution to translation education is also an improvement embedded in the data-driven models like Transformer-GAN to observe what is going on to the entire classroom and modify the content of his/her curriculum accordingly. Such models incorporated in learning management systems (LMS) would make translation training efficient, scalable, and individualized.

Thus, the future holds inordinate prospects for AI enriching the learning surface as it stands by propagating student-centered learning environments, while harboring inquiry toward the development of intelligent translation tools as a significant part of modern language education.

4. Conclusion

The study was intended to bring into the use of Transformer-GAN based model useful evaluation and enhancement of English translation skill especially college students. The model made such a perfect combination between the contextual powerful capabilities of the Transformers and that of the refinement afforded by the GANs to achieve superior results in translation accuracy, fluency, and cultural appropriateness. Compared to the conventional models and BiLSTM + FIS, all of the vital metrics recorded have shown that Transformer-GAN performed better than any of them in diagnosis and correction of different types of translation errors. It also affords a form of automated individualized feedback to both teaching and self-study.

In short, Transformer-GAN will be a powerful, intelligent solution toward contemporary translation education, thereby advocating improved learning as well as supporting teachers and easily being adopted in future classes with AI integrated.

References

[1] Kruk, M., & Kałużna, A. (2025). Investigating the Role of AI Tools in Enhancing Translation Skills, Emotional Experiences, and Motivation in L2 Learning. *European Journal of Education*, *60*(1), e12859.

[2] Tutton, M., & Cohen, D. (2025). Reconceptualizing the Role of the University Language Teacher in Light of Generative AI. *Education Sciences*, *15*(1), 56.

[3] Minas, D., Theodosiou, E., Roumpas, K., & Xenos, M. (2025). Adaptive Real-Time Translation Assistance Through Eye-Tracking. *AI*, *6*(1), 5.

[4] Zhang, N., & He, X. Z. (2024). Corpus-based research on English-Chinese translation teaching combining vocabulary learning and practice. *Journal of Electrical Systems*, 20(3SI), 2855-2866. https://doi.org/10.52783/jes.3187.

[5] Zhang, L. (2022). Research on English translation teaching model: based on key process area model. *Journal of Environmental Protection and Ecology*, *23*(4), 1693-1699.

[6] Lai, N. (2024). An IoT-based multiple teaching quality evaluation method for English translation with improved deep learning. *Engineering Reports*, *6*(11), e12896.

[7] Zhang, P. X. (2024). Design and implementation of English-Chinese translation teaching platform based on deep learning. *Journal of Electrical Systems*, 20(3SI), 1746-1755. https://doi.org/10.52783/jes.1714.

[8] Qiao, H., An, L., & Cao, M. (2023). RETRACTED ARTICLE: Development of English translation online teaching system and service function improvement based on image segmentation algorithm. *Soft Computing*, *27*(12), 8415-8424.

[9] Bin, W. (2023). Application of improved image restoration algorithm and depth generation in English intelligent translation teaching system. *Soft Computing*, 1-11.

[10] Kong, B., & He, C. (2025). Deep Learning and Fuzzy Algorithm in Improving the Effectiveness of College English Translation Teaching. *Computers and Education: Artificial Intelligence*, 100378.

[11] Boadu, S. K., & Bannor, G. A. (2023). Investigating Hindrance Factors Which Forestalled Teachers'participation in Performance Contracts In Senior High Schools In Ghana. *Matrix Science Mathematic*, *7*(2), 88-94.

[12] Whendasmoro, R. G., & Joseph, J. (2022). Analisis Penerapan Normalisasi Data Dengan Menggunakan Z-Score Pada Kinerja Algoritma K-NN. JURIKOM (Jurnal Riset Komputer), 9 (4), 872Seo, H., Hwang, T., Jung, J., Kang, H., Namgoong, H., Lee, Y., & Jung, S. (2025). Large Language Models as Evaluators in Education: Verification of Feedback Consistency and Accuracy. *Applied Sciences (2076-3417)*, *15*(2).

[13] Duan, G., Zheng, X., Zhu, Y., Ren, T., & Yan, Y. (2023, May). An Efficient Transformer with Distance-aware Attention. In *2023 IEEE 9th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing,(HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)* (pp. 96-101). IEEE.

[14] Liu, H., Hou, R., & Lepage, Y. (2024). High-Quality Data Augmentation for Low-Resource NMT: Combining a Translation Memory, a GAN Generator, and Filtering. *arXiv preprint arXiv:2408.12079*.

[15] Wu, A. N., Stouffs, R., & Biljecki, F. (2022). Generative Adversarial Networks in the built environment: A comprehensive review of the application of GANs across data types and scales. *Building and Environment*, *223*, 109477.