

## PAPER

# Modeling the Effectiveness of an AR-Based Context-Aware Platform for English Translation Learning

Xuejiao Liu<sup>1</sup> ,  
Yinghua Hu<sup>2</sup>  (✉),  
Zhiyun Zhang<sup>2</sup> 

<sup>1</sup>Hebei University of  
Environmental Engineering,  
Qinhuangdao, China

<sup>2</sup>Yanshan University,  
Qinhuangdao, China

[hu124yinghua@163.com](mailto:hu124yinghua@163.com)

## ABSTRACT

Traditional English translation instruction has long faced persistent challenges, including the absence of authentic contexts, delayed feedback, and insufficient personalization. With the advancement of educational technologies, mobile learning platforms that integrate augmented reality (AR) and context-aware mechanisms have opened new pathways for creating immersive and situational translation learning environments. However, existing studies have primarily concentrated on technological implementation and user satisfaction, while providing limited quantitative evidence regarding the key factors that determine instructional effectiveness. In response to this gap, a modeling-based investigation was conducted to examine the mechanisms through which an integrated AR and context-aware mobile interactive platform influences translation learning outcomes. A prototype platform was designed and developed to deliver visual contextual support through AR and to dynamically adapt learning content via context-aware sensing. Through instructional experiments, pre- and post-test translation performance data, platform interaction logs, and contextual information were systematically collected. A random forest regression model was employed, with translation competence improvement as the predictive target, to conduct multidimensional feature modeling and analysis. The results demonstrated that the experimental group achieved significantly greater improvements in overall translation competence, as well as in linguistic accuracy, pragmatic appropriateness, and contextual adaptability, compared to the control group, confirming the instructional effectiveness of the AR-based context-aware platform. Moreover, the random forest model exhibited superior predictive accuracy compared with traditional models. AR scene interaction frequency and contextual complexity were found to be the primary driving features, showing a positive interaction effect—enhanced benefits were observed under conditions of high contextual complexity combined with deep AR interactions.

## KEYWORDS

augmented reality (AR), context awareness, English translation instruction, random forest regression, educational data mining, feature importance, mobile interactive platform

Liu, X., Hu, Y., Zhang, Z. (2025). Modeling the Effectiveness of an AR-Based Context-Aware Platform for English Translation Learning. *International Journal of Interactive Mobile Technologies (IJIM)*, 19(24), pp. 45–59. <https://doi.org/10.3991/ijim.v19i24.59473>

Article submitted 2025-08-30. Revision uploaded 2025-10-25. Final acceptance 2025-11-03.

© 2025 by the authors of this article. Published under CC-BY.

## 1 INTRODUCTION

In the era of globalization, English translation competence has become one of the core literacies essential for cross-cultural communication and professional collaboration [1, 2]. However, traditional English translation instruction has long been constrained by several structural challenges, which have hindered the deep enhancement of instructional effectiveness. Foremost among these challenges is the absence of authentic contexts [3, 4]. Conventional classrooms often isolate language from its original cultural and social settings, making it difficult for learners to achieve precise meaning transfer and functional equivalence when translating real-world scenarios such as museum tours or restaurant dialogues [5]. Second, instructional methods remain overly abstract and monotonous, relying excessively on text substitution and grammar-based explanations, which fail to stimulate learners' interest or foster deep cognitive engagement [6, 7]. Furthermore, feedback is typically delayed, as instructors are unable to provide immediate and individualized feedback on students' translation practice, resulting in the reinforcement of persistent errors [8]. Finally, instructional personalization is inadequate, with existing pedagogical models struggling to accommodate learners' diverse learning styles and proficiency levels [9].

Concurrently, the rapid advancement of information technology has infused educational transformation with renewed momentum. The widespread adoption of mobile learning has broken the spatial and temporal constraints of traditional education, enabling “ubiquitous learning” across multiple environments [10, 11]. Among emerging technological integrations, the fusion of AR and context-aware mechanisms has demonstrated remarkable potential in addressing the aforementioned pedagogical challenges. AR technology enables the seamless overlay of virtual information onto the real world, constructing an immersive visual context that directly links abstract linguistic symbols with tangible visual referents [12, 13]. Meanwhile, context-aware mechanisms can capture learners' physical environments, temporal conditions, and behavioral states, thereby delivering real-time, contextually relevant learning content. This capability facilitates a paradigm shift from uniform instruction to context-responsive learning, in which instructional delivery dynamically adapts to situational variables [14, 15].

Although preliminary attempts have been made to incorporate AR and context-aware technologies into language instruction, a significant research gap remains. The majority of existing studies have concentrated on technical development and platform implementation or have relied primarily on questionnaires and interviews to describe learners' subjective experiences such as satisfaction and motivation [16, 17]. While these findings are valuable, they fail to address two essential questions: (a) How can the actual effectiveness of such integrated technologies on translation competence—the core instructional objective—be quantitatively assessed? (b) Within the vast interactional and contextual data generated by such platforms, which specific factors fundamentally drive learning outcomes? The lack of clarity surrounding these mechanisms has obscured optimization directions for educational technology developers and deprived frontline instructors of precise data-driven support for instructional design.

The overall framework of the study followed a “platform construction–data collection–modeling analysis” sequence. First, a prototype of a mobile interactive English translation instruction platform integrating AR and context-aware mechanisms was designed and developed. Subsequently, controlled instructional experiments were conducted to systematically collect learners' behavioral interaction data,

contextual information, and pre- and post-test translation competence scores. Finally, a random forest regression algorithm was introduced, with translation performance improvement used as the dependent variable and multidimensional platform interaction features as independent variables. An instructional effectiveness prediction model was thus constructed, and feature importance ranking together with partial dependence analysis was employed to explore the intrinsic mechanisms underlying the observed learning outcomes.

## 2 RESEARCH DESIGN AND PLATFORM DEVELOPMENT

### 2.1 Design of the mobile interactive platform integrating AR and context-aware mechanisms

The mobile interactive platform developed in this study was designed around a “perception–comprehension–augmentation” closed-loop system, intended to computationally and operationally implement SLT and to address the long-standing disconnection between linguistic symbols and authentic contexts in traditional translation instruction. The platform architecture was developed in accordance with the principle of separation of concerns, adopting a three-tier structure comprising the client side, server side, and data layer. Computational tasks were strategically distributed across the layers to balance real-time interaction performance on mobile devices with the demands of complex data analysis. The client side integrates computer vision algorithms, multimodal sensors, and a graphical user interface to perform real-time acquisition of environmental and user behavioral data, as well as initial AR rendering. The server side, functioning as the intelligent hub, stores the AR content library and context-rule repository, and embeds a natural language processing (NLP) engine responsible for executing complex tasks such as image recognition, context inference, and translation evaluation. The data layer employs a structured storage framework to persistently record interaction logs, user profiles, and learning content, thus establishing a traceable quantitative data foundation for subsequent modeling.

### 2.2 Experimental design

To rigorously verify the instructional effectiveness of the constructed platform, a pre-test–intervention–post-test design with a control group was adopted. The design followed the principles of controlled comparison and causal inference, ensuring that any differences in learning outcomes could be attributed to the platform intervention rather than to confounding variables. The core logic of the experimental procedure was to establish a baseline of participants’ initial translation competence through a pre-test, which was then used as a covariate to control for individual differences. During the intervention phase, the experimental group received instruction via the AR context-aware mobile interactive platform for eight weeks (two sessions per week, 90 minutes per session), while the control group underwent an equivalent duration and intensity of traditional mobile-based translation instruction. The post-test, structurally and in terms of difficulty identical to the pre-test, was administered immediately after the intervention to measure changes in translation performance. To enhance internal validity, a random cluster sampling method was used to assign

participants to the experimental and control groups. An independent samples t-test confirmed that no significant difference existed between the two groups' pre-test scores ( $p > 0.05$ ). Multiple control measures were further implemented: (a) identical instructional procedures and materials were used across both groups to eliminate instructor bias; (b) external learning resources were restricted during the intervention to minimize history effects; and (c) follow-up interviews were conducted to reduce attrition bias. These measures collectively mitigated potential threats to validity, including maturation and testing effects.

The variable definition system adhered to the core logic of “theoretical construct–operational definition–measurement method–data source.” It focused on the three core interactive dimensions of learner–technology–context, selectively retaining the key indicators that most effectively captured the underlying mechanisms. This design ensured both theoretical depth and empirical rigor while maintaining the feasibility of data acquisition and the compatibility of the model. The framework was presented through an integrated combination of explanatory text and standardized tables, clarifying the underlying logic of variable design and presenting the detailed indicators in a concise and structured manner.

**Table 1.** Design of dependent variable indicator

Feature Name	Operational Definition	Data Type	Value Range	Data Source
Translation competence gain score	The difference between the post- and pre-test total translation scores. Evaluation dimensions include linguistic accuracy (40%), pragmatic appropriateness (30%), and contextual adaptability (30%). Scores were assigned through double-blind evaluation by two senior instructors (Intra-class correlation coefficient, ICC = 0.87).	Continuous	$(-\infty, +\infty)$	Standardized test responses from the pre-test (one week before intervention) and post-test (immediately after intervention)

The dependent variable focused on the net improvement in translation competence, representing the learner's dynamic developmental trajectory during the experimental intervention rather than absolute performance levels. A gain-score model was employed to quantify this improvement, effectively controlling for individual differences in initial competence and more precisely reflecting the net effect of the instructional intervention. The assessment of translation competence was based on Bachman's Communicative Language Ability (CLA) framework and the Translation Competence Framework for Translation Studies, forming a three-dimensional evaluation system encompassing linguistic accuracy, pragmatic appropriateness, and contextual adaptability. The standardized assessment instrument underwent two rounds of expert validation (Content Validity Index, CVI = 0.92) and pilot testing (Cronbach's  $\alpha = 0.89$ ), demonstrating strong reliability and validity. The specific design of the dependent variable indicator is presented in Table 1.

The system of independent (feature) variables was established based on a “learner–technology–context” interaction framework, subdivided into four major categories. Each category was closely anchored to core theoretical constructs, with refined indicators designed to capture key influencing factors while avoiding redundant information that could weaken the interpretability of the model.

- a) Platform usage behavior features: This category of variables was grounded in TAM and the Self-Regulated Learning (SRL) Theory. Its primary objective was to quantify the interactional processes between learners and the platform into computable indicators that reflect cognitive strategies and learning engagement.

By selecting only those behavioral metrics directly associated with learning outcomes and excluding secondary operational records, the indicators were optimized to accurately capture the intrinsic logic of active participation, strategic application, and outcome transformation. The core indicators are presented in Table 2.

**Table 2.** Design of platform usage behavior feature indicators

Feature Name	Operational Definition	Data Type	Value Range	Data Source
AR scene interaction frequency	Grouped by tertiles based on experimental data: low frequency (< 1.2 interactions/min, lower tertile), medium frequency (1.2–2.5 interactions/min, middle tertile), and high frequency (> 2.5 interactions/min, upper tertile). Used for subsequent interaction effect analyses.	Continuous	[0, +∞)	Client-side interaction logs
Feedback viewing time ratio	Ratio of cumulative feedback-viewing time to total task duration	Continuous	[0, 1]	Client-side behavioral sequence logs
Task completion rate	Proportion of completed tasks relative to the total number of system-assigned tasks	Continuous	[0, 1]	Server-side task management system
Self-initiated task ratio	Ratio of voluntarily selected non-system-assigned tasks to total tasks	Continuous	[0, 1]	Server-side task type logs
Multimodal input switching frequency	Number of mode switches per unit time among text input, voice input, and gesture annotation	Continuous	[0, +∞)	Client-side input mode logs

**b)** Contextual features: The contextual features were based on the three-dimensional situational model of “physical environment–social interaction–task objective” proposed in SLT. Its core aim was to transform the dynamic context perceived by the AR platform into quantifiable key indicators. The selection criteria focused on attributes that were directly relevant to translation learning and reflective of contextual adaptability, while irrelevant scene details were excluded to ensure that the indicators effectively captured the underlying associative mechanism among context, task, and learning outcome. The core indicators are presented in Table 3.

**Table 3.** Design of contextual feature indicators

Feature Name	Operational Definition	Data Type	Value Range/Classification	Data Source
Context type	Classification of physical scenes based on GPS and computer vision recognition: (a) daily-life scenario, (b) educational scenario, (c) professional scenario, (d) public service scenario	Categorical	1–4	GPS + computer vision recognition logs
Context complexity	Composite score derived from the number of scene elements and task objectives: 1 = single-object translation, 2–4 = moderate complexity, 5 = multi-role interactive translation	Ordinal	1–5	Server-side contextual evaluation module
Task relevance	Degree of alignment between the current task and the learner’s professional domain, calculated based on pre-registered discipline tags	Continuous	[0, 1]	Server-side user model + task tags
Social participation level	Number of participants involved in the translation task: 1 = single-user, 2 = pair, 3 = group	Categorical	1–3	Camera-based facial recognition + user annotations

## c) AR interaction quality features

**Table 4.** Design of AR interaction quality feature indicators

Feature Name	Operational Definition	Data Type	Value Range	Data Source
AR registration accuracy	Average deviation between the virtual information anchor point and the corresponding physical object (<10 cm = optimal; >30 cm = poor)	Continuous	[0, +∞) (cm)	Computer vision localization logs
Rendering latency	Average response time from scene change to AR content update (<100 ms = optimal; >500 ms = poor)	Continuous	[0, +∞) (ms)	Client-side rendering engine logs
Interaction success rate	Ratio of successfully recognized user operations (e.g., taps and voice commands) to total operation attempts	Continuous	[0, 1]	Client-side input recognition logs
System stability	Number of system crashes or freezes per unit of learning duration	Count	[0, +∞)	Server-side error monitoring logs

This category of features was established in accordance with the ISO 9241–11 standard for human–computer interaction usability, with the core objective of extracting technical performance indicators that directly influence both learning experience and instructional effectiveness. The selection principles were guided by objectivity, measurability, and high contextual relevance to AR-based learning environments. Technical parameters unrelated to the learning process were excluded to ensure that the selected indicators accurately reflected the transmission pathway among technical performance, interaction experience, and learning outcome. The core indicators are presented in Table 4.

- d) Learner individual features: As control variables, this category of features was established based on Differentiated Learning Theory and TAM. The core objective was to isolate the influence of individual differences on learning outcomes and to focus on learner attributes that are strongly associated with translation learning and AR technology usage. The selection criteria emphasized variables that significantly affect learning engagement and technological adaptability, ensuring that these indicators effectively controlled for confounding variables and thereby clarified the causal relationship between platform interaction and learning effectiveness. The core indicators are presented in Table 5.

**Table 5.** Design of learner individual feature indicators

Feature Name	Operational Definition	Data Type	Value Range/ Classification	Data Source
Initial English proficiency	Results of the Oxford English Placement Test (OEPT) based on the CEFR framework, categorized into six proficiency levels (A1-C2)	Ordinal	A1-C2	Pre-intervention English proficiency test
Translation learning experience	Cumulative duration of formal translation learning, including coursework, training, and practice activities	Continuous	[0, +∞)	Learner background questionnaire
Technology acceptance	TAM-based scale score (7-point scale), comprising three dimensions: perceived usefulness, perceived ease of use, and behavioral intention to use	Continuous	[3, 21]	Technology acceptance questionnaire ( $\alpha = 0.85$ )
Learning motivation	Intrinsic motivation score derived from the AMOS scale (5-point scale), including key dimensions such as interest and self-efficacy	Continuous	[5, 25]	Learning motivation questionnaire ( $\alpha = 0.82$ )

The integrated variable framework was articulated through explanatory paragraphs that clarified the theoretical foundation and design logic of each variable category, while standardized tables were employed to present the operational details of the indicators in a structured manner. This dual representation achieved an organic integration of “theoretical depth” and “practical operability.” In terms of construct validity, all indicators were tightly aligned with established theoretical frameworks and were reviewed by domain experts to ensure consistency with the underlying constructs. Regarding data adaptability, the system incorporated a hybrid structure encompassing continuous, categorical, and ordinal variables, allowing for optimal compatibility with the random forest model. Standardization procedures were applied to eliminate dimensional bias. To prevent model overfitting, redundancy was minimized through correlation analysis, retaining only the most influential indicators.

### 3 EFFECTIVENESS MODELING BASED ON RANDOM FOREST REGRESSION

#### 3.1 Model construction and training

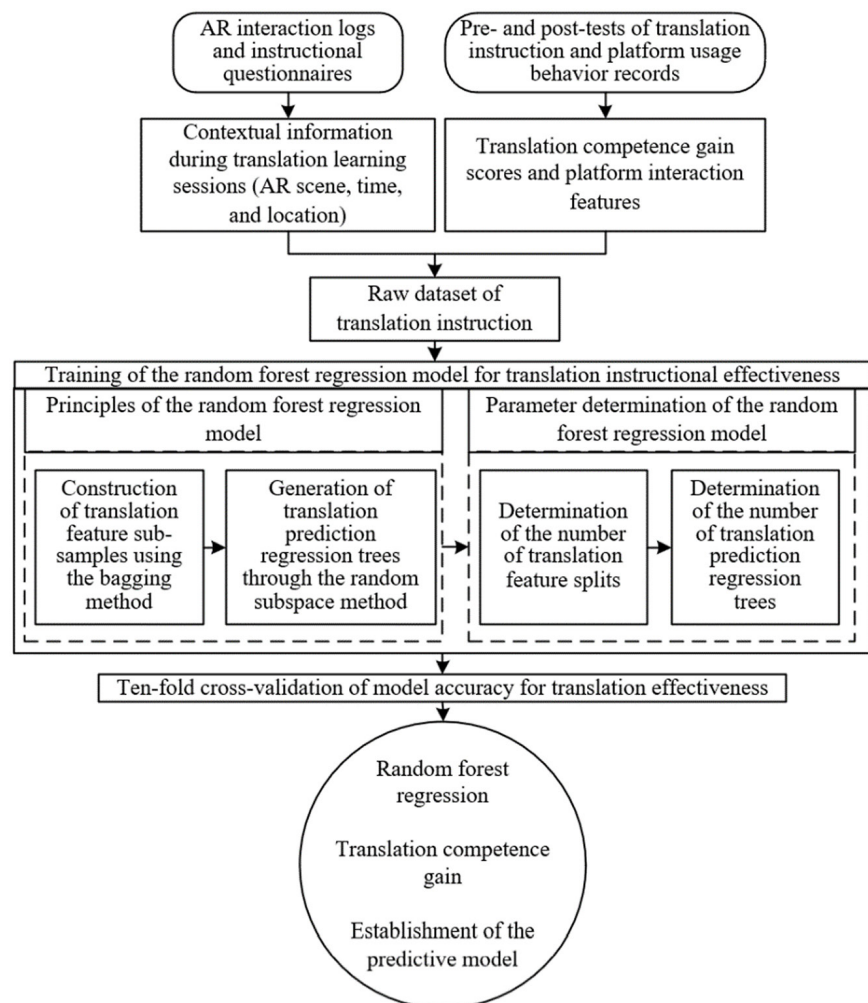


Fig. 1. Framework for effectiveness modeling of AR and context-aware translation instruction based on random forest regression

The construction and training of the model were conducted following the standardized machine learning workflow of data partitioning, hyperparameter optimization, model fitting, and performance validation. The core objective was to accurately capture the intrinsic relationship between feature variables and translation competence gain while ensuring strong model generalization capability. A holdout validation method was employed to randomly partition the preprocessed dataset into a training set and a testing set at a ratio of 7:3, with a fixed random seed applied to guarantee reproducibility. The choice of this partition ratio was based on statistical reasoning: a 70% training set provides sufficient data for the random forest model to learn complex relationships, whereas a 30% testing set ensures a stable estimation of generalization error. This ratio aligns with established best practices in machine learning research. The training set was utilized for model parameter fitting and hyperparameter tuning. The testing set, treated as unseen data, served as an independent dataset to evaluate the model's predictive performance on new samples. The framework for effectiveness modeling of AR and context-aware translation instruction based on random forest regression is illustrated in Figure 1.

The hyperparameters directly determined the complexity and learning capacity of the model. The essence of optimization was to identify the optimal equilibrium point within the bias–variance trade-off: an overly simple model tends to underfit the data, while an excessively complex one risks overfitting, fitting noise rather than the underlying structure. A combined strategy of grid search and tenfold cross-validation was employed for hyperparameter optimization. Based on both literature review and domain expertise, the search space for key hyperparameters was defined as follows: number of decision trees ( $n_{estimators}$ ): [100, 200, 300, 500]; maximum tree depth ( $max_{depth}$ ): [None, 5, 10, 20]; minimum samples required to split a node ( $min_{samples\_split}$ ): [2, 5, 10]. During training, tenfold cross-validation was conducted within the training set. The training set was divided into 10 mutually exclusive subsets, where nine subsets were used for training and one for validation in each iteration. The mean coefficient of determination ( $R^2$ ) across the ten validation rounds was computed as the performance metric, and the hyperparameter configuration achieving the highest average  $R^2$  was selected as the optimal parameter combination.

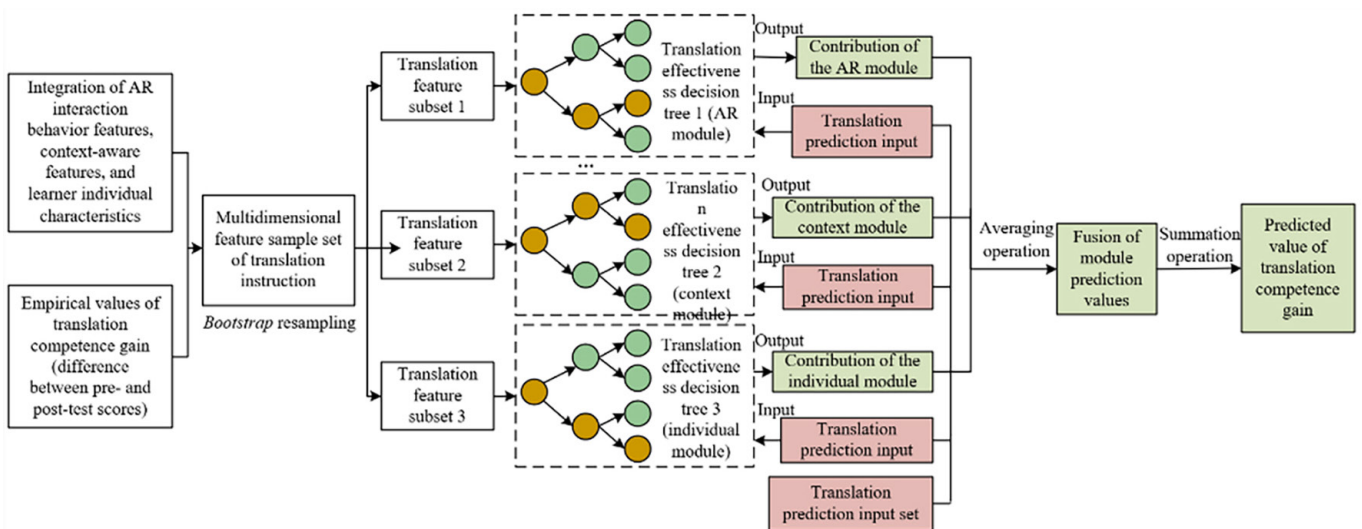


Fig. 2. Feature–effect modeling pathway for AR and context-aware translation instruction based on random forest regression

The final model was fitted using the optimal hyperparameters and the training dataset. The fitting process deeply integrated the two core principles of ensemble learning—bootstrap aggregation (bagging) and the random subspace method. The mechanism operated below. First, using bootstrap sampling, data subsets equal in size to the original training set were drawn with replacement to construct individual decision trees. Each tree was trained on approximately 63.2% unique samples, while the remaining 36.8% of unsampled data constituted the out-of-bag (OOB) samples, which were used for internal model performance evaluation without requiring an additional validation set. Second, at each node split within a tree, instead of selecting the optimal split from all available features, a random subset comprising one-third of the total feature set was drawn as candidate variables, following the standard practice and feature dimension in random forest regression. This approach balanced model diversity and learning efficiency. The dual-randomization mechanism was designed to generate a set of diverse and mildly underfitted base learners. Each tree learned from distinct perspectives of the data and feature dimensions, resulting in prediction errors that were largely independent of one another. Finally, by averaging the regression outputs of all decision trees, individual errors were effectively offset, yielding a model characterized by low bias, reduced variance, and robust generalization capability.

As shown in Figure 2, the multi-module modeling process for AR and context-aware translation instruction based on random forest regression was established through the systematic integration of multi-source features—including AR interaction behavior, context awareness, and individual learner characteristics—with the empirically observed translation competence gain, thereby constructing a sample set. Using bootstrap resampling, multiple feature subsets were generated to construct decision trees for each module. Each module independently produced its contribution output. The module-level outputs were subsequently fused to obtain the final predicted translation competence gain. This two-stage process, characterized by modular modeling and contribution fusion, simultaneously leveraged the advantages of ensemble learning inherent in random forest and enabled the precise decomposition of the effects of AR, context, and learner-specific attributes on translation instructional effectiveness.

### 3.2 Model evaluation and result analysis

To quantitatively assess the accuracy differences between the AR and context-aware random forest regression model and traditional regression models in predicting translation competence gain, and to verify the superiority of the proposed modeling approach, three representative models were compared across core performance metrics, as shown in Table 6. The results indicate that  $R^2$  (0.87) of the random forest regression model was substantially higher than those of multiple linear regression (0.62) and SVR (0.71). This demonstrates that 87% of the variance in translation competence gain was explained by the random forest model, evidencing a markedly higher explanatory capability than traditional models. RMSE (2.15) and MAE (1.62) of the random forest model were the lowest among all models, showing reductions of 44.7% and 44.9% compared with multiple linear regression and 34.1% and 35.5% compared with SVR, respectively. The MAPE (7.8%) was only 54.5% of that of multiple linear regression (14.3%) and 67.8% of that of SVR (11.5%), reflecting substantially higher prediction stability. These findings clearly demonstrate that the random forest regression model integrating multidimensional data,

such as AR interaction behavior and contextual features, accurately captured the complex nonlinear relationships underlying translation instructional effectiveness. In terms of predictive accuracy, explanatory power, and robustness, it significantly outperformed traditional regression models. The results thus provide strong empirical support for the AR and context-aware mechanism combined with random forest modeling as a highly effective and adaptive framework for the quantitative analysis of translation instructional effectiveness.

To identify the core factors driving translation competence gain, feature contributions were quantified using both Gini importance and permutation importance derived from the random forest model. As shown in Table 7, AR scene interaction frequency, proportion of feedback viewing duration, and contextual complexity ranked highest among all features. Notably, features related to AR interaction and feedback exhibited substantially higher importance scores than individual learner features (e.g., initial English competence and learning motivation) and technical parameters (e.g., AR registration accuracy). These findings indicate that the depth of learner-AR platform interaction and the complexity of learning contexts play decisive roles in promoting translation competence gain. This observation empirically validates the research hypothesis that the AR and context-aware mechanism enhance learning outcomes by strengthening interactivity and contextual adaptability. Furthermore, it provides a data-driven foundation for optimizing platform functionality and refining instructional strategies in translation education.

**Table 6.** Experimental results of quantitative model performance indicators

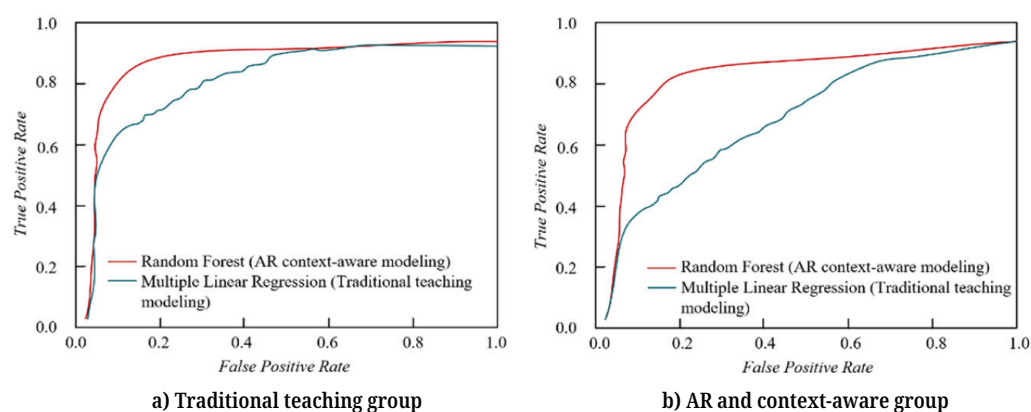
Model Type	R <sup>2</sup>	Root Mean Square Error (RMSE)	Mean Absolute Error (MAE)	Mean Absolute Percentage Error (MAPE, Added)
Random forest regression (AR and context-aware)	0.87	2.15	1.62	7.8%
Multiple linear regression (traditional)	0.62	3.89	2.94	14.3%
Support Vector Regression (SVR, traditional)	0.71	3.26	2.51	11.5%

**Table 7.** Experimental results of feature importance ranking

Feature Name	Gini Importance Score (Normalized)	Permutation Importance Score (Normalized)
AR scene interaction frequency	0.18	0.16
Proportion of feedback viewing duration	0.15	0.14
Contextual complexity	0.12	0.13
Task relevance	0.10	0.11
Multimodal input switching frequency	0.08	0.09
Initial English competence	0.07	0.08
AR registration accuracy	0.06	0.07
Learning motivation	0.05	0.06

To evaluate the model's capability in identifying cases of significant improvement in translation competence and to further validate the value of the AR and context-aware mechanism in precisely locating high-performing learners, a binary classification label was defined based on the criterion of gain score  $\geq$  mean + 1

standard deviation (SD). The receiver operating characteristic (ROC) curves and area under the curve (AUC) values of different models were then compared. As illustrated in Figure 3, in the traditional teaching group, the random forest model achieved an AUC of 0.90, markedly higher than the 0.72 of multiple linear regression, indicating superior identification accuracy for high-performing learners within conventional instructional settings. In the AR and context-aware group, the AUC of the random forest model further increased to 0.95, widening the performance gap from 0.72 to 0.95 compared with multiple linear regression. This suggests that after integrating AR interaction and contextual features, the proposed model not only achieved precise continuous prediction of translation competence gain but also efficiently identified learners with significant improvement, exhibiting classification superiority over traditional models. The practical implication of this finding lies in its potential to support personalized instructional interventions. Learners predicted as “non-significantly improved” can be proactively identified, allowing targeted optimization of AR platform interaction design or instructional strategies, further enhancing instructional effectiveness. From a binary classification perspective, these results also reaffirm the supporting role of the AR and context-aware mechanism in achieving the core instructional objectives of translation teaching.



**Fig. 3.** ROC curves of random forest regression and multiple linear regression models in translation instruction

To directly validate the instructional effectiveness of the AR and context-aware mobile interactive platform, an independent samples t-test was conducted to compare the learning outcomes between the experimental and control groups. As shown in Table 8, no significant differences were observed in any pre-test indicators between the two groups ( $p > 0.05$  and  $d \leq 0.06$ ), confirming the baseline equivalence and fairness of the experimental intervention. Following the intervention, the post-test total score of the experimental group ( $82.67 \pm 6.43$ ) was significantly higher than that of the control group ( $70.12 \pm 7.28$ ). The translation competence gain score reached  $20.32 \pm 5.17$  in the experimental group—2.47 times higher than the  $8.23 \pm 4.69$  of the control group. Significant differences ( $p < 0.001$ ) were also observed across all post-test and gain indicators. From the perspective of effect size, the total score, language accuracy, and pragmatic appropriateness all exhibited large effects (Cohen’s  $d > 1.5$ ), while contextual adaptability showed a moderate-to-large effect ( $d = 0.93$ ), indicating substantial practical significance of the differences. These results confirm that the AR and context-aware platform exert a pronounced instructional advantage. By integrating contextual immersion, multimodal interaction, and formative feedback

into a closed-loop design, the platform effectively enhanced learners' translation competence. The findings strongly support the core research hypothesis that the AR and context-aware mechanism significantly promote the instructional effectiveness of English translation teaching.

**Table 8.** Experimental results of learning outcome differences between the experimental and control groups

Measurement Indicator	Experimental Group (AR and Context-Aware Platform, n = 60)	Control Group (Traditional Teaching, n = 60)	t-Value	p-Value	Effect Size (Cohen's d)
Pre-test total score (100)	62.35 ± 7.82	61.89 ± 8.15	0.32	0.751	0.05
Post-test total score (100)	82.67 ± 6.43	70.12 ± 7.28	9.87	<0.001	1.56
Translation competence gain score	20.32 ± 5.17	8.23 ± 4.69	12.45	<0.001	2.01
Language accuracy pre-test (40)	24.92 ± 3.51	24.76 ± 3.68	0.25	0.803	0.04
Language accuracy post-test (40)	34.68 ± 2.87	28.95 ± 3.21	10.12	<0.001	1.62
Language accuracy gain score	9.76 ± 2.13	4.19 ± 1.87	13.06	<0.001	2.11
Pragmatic appropriateness pre-test (30)	18.71 ± 2.64	18.53 ± 2.79	0.37	0.712	0.06
Pragmatic appropriateness post-test (30)	25.83 ± 2.15	21.46 ± 2.48	9.54	<0.001	1.52
Pragmatic appropriateness gain score	7.12 ± 1.78	2.93 ± 1.62	12.89	<0.001	2.08
Contextual adaptability pre-test (30)	18.72 ± 2.75	18.60 ± 2.81	0.23	0.818	0.04
Contextual adaptability post-test (30)	22.16 ± 2.03	19.71 ± 2.35	5.78	<0.001	0.93
Contextual adaptability gain score	3.44 ± 1.36	1.11 ± 1.29	9.67	<0.001	1.55

Notes: Cohen's  $d \geq 0.8$  indicates a large effect, 0.5–0.8 a medium effect, and  $<0.5$  a small effect.

**Table 9.** Experimental results of the interaction effects between core features

AR Scene Interaction Frequency (Tertile Grouping)	Low Contextual Complexity	Medium Contextual Complexity	High Contextual Complexity	Marginal Gain of Interaction Effect (Medium → High Context)
Low (< 1.2 times/min)	12.67 ± 3.21	15.83 ± 3.64	16.92 ± 3.87	1.09
Medium (1.2–2.5 times/min)	16.89 ± 3.58	20.51 ± 4.12	25.36 ± 4.35	4.85
High (> 2.5 times/min)	18.72 ± 3.74	23.45 ± 4.28	29.81 ± 4.56	6.36
Marginal gain of main effect (low → high interaction)	6.05	7.62	12.89	–

To investigate how the interaction between core features influences translation competence gain, a partial dependence analysis combined with grouped statistics was conducted, focusing on the interaction between AR scene interaction frequency and contextual complexity. The results presented in Table 9 indicate a significant positive interaction effect between the two variables. Under conditions of low contextual complexity, the marginal gain of the main effect of increasing AR scene interaction frequency from low to high was only 6.05, indicating limited improvement. In contrast, under high contextual complexity, the corresponding marginal gain reached 12.89, which is 2.13 times greater than that observed in low-complexity contexts. Moreover, the marginal gain of the interaction effect exhibited a monotonic increase as AR interaction frequency rose. When the AR interaction frequency was low, the marginal gain from medium to high contextual complexity was merely

1.09, whereas under high interaction frequency, it increased to 6.36. This pattern demonstrates that in high-complexity contexts, learners who engage more intensively in AR-based interaction can more effectively leverage contextual cues, thereby deepening their comprehension of contextualized translation rules. This finding reveals a synergistic amplification mechanism between AR interaction depth and contextual complexity, providing a precise basis for differentiated instructional design. AR interaction functionality should be reinforced in high-complexity contexts, whereas in low-complexity contexts, instructional focus may be better placed on foundational translation training, further substantiating the core conclusion that contextualization and interactivity constitute the essential logic through which AR platforms enhance the instructional effectiveness of translation teaching.

## 4 CONCLUSION

The AR and context-aware interactive platform with a “perception–comprehension–enhancement” closed-loop was constructed, and its pedagogical effectiveness and underlying mechanisms in translation teaching were investigated through a “pre-test–intervention–post-test with control group” experimental design combined with random forest regression modeling. The findings demonstrated that the experimental group achieved significantly higher total translation competence scores and three-dimensional gain scores compared with the control group ( $p < 0.001$  and effect size = 1.52–2.11), confirming the practical effectiveness of the platform. The random forest model integrating multidimensional features exhibited optimal predictive performance ( $R^2 = 0.87$ , RMSE = 2.15, and AUC = 0.95), outperforming traditional regression models. Among all predictors, AR scene interaction frequency, feedback viewing duration ratio, and contextual complexity were identified as the core driving features, and a synergistic amplification mechanism was observed between AR interaction depth and contextual complexity. The principal contributions of this study are threefold. Theoretically, it computationally operationalized SLT, thereby enriching the integration framework of educational technology and translation pedagogy. Methodologically, it validated the superiority of random forest algorithms in processing high-dimensional, nonlinear educational data and established a new paradigm for instructional effectiveness modeling. Practically, it proposed an implementable solution to the persistent challenges of contextual disconnection and insufficient interaction in traditional translation instruction.

Certain limitations remain. The sample was limited to university learners within a specific region, constraining both group diversity and sample size. The eight-week intervention did not track long-term effects or the transferability of translation competence. Furthermore, the feature system did not incorporate affective or emotional state variables, and the AR technology was limited in compatibility with mainstream devices. Future research will expand the sample scope and scale, extend the intervention duration, and include follow-up testing. The feature system will be enriched through the integration of multisource data. Concurrently, the platform’s cross-device adaptability and immersive AR experience will be further optimized.

## 5 ACKNOWLEDGEMENTS

This paper was funded by the 2025 Hebei Province Regular Undergraduate Colleges English Teaching Reform Research and Practice Project titled “Research on

the ‘Entrepreneurship Simulation’ Teaching Model of College English Translation under the Synergy of Human, Data and AI” (Grant No.: 2025YYJG012); and the Chinese Academic Translation Project of the National Social Science Fund titled “Governance Transformation and NGO Development in Contemporary China (English Version)” (Grant No.: 23WSHB009).

## 6 REFERENCES

- [1] X. Zhao and Y. Jiang, “Synchronously improving multi-user English translation ability by using AI,” *International Journal on Artificial Intelligence Tools*, vol. 31, no. 4, p. 2240007, 2022. <https://doi.org/10.1142/S0218213022400073>
- [2] Y. Li, M. Huang, and H. Dong, “Interactive mobile English translation proficiency model based on particle swarm optimisation and neural network for teaching,” *International Journal of Interactive Mobile Technologies*, vol. 19, no. 8, pp. 109–124, 2025. <https://doi.org/10.3991/ijim.v19i08.54043>
- [3] S. V. Resende, “Interculturality in English teaching: A translation-based perspective,” *Humanidades & Inovacao*, vol. 5, no. 7, pp. 198–211, 2018.
- [4] H. Chang and J. Zhao, “Enhancing translation teaching efficiency and learning experience through mobile technology applications,” *International Journal of Interactive Mobile Technologies*, vol. 19, no. 6, pp. 75–89, 2025. <https://doi.org/10.3991/ijim.v19i06.54703>
- [5] K. Sun, “Teaching English-Chinese textual translation strategies: A topic-chain approach,” *Babel*, vol. 65, no. 2, pp. 286–315, 2019. <https://doi.org/10.1075/babel.00097.sun>
- [6] L. Cherchata, L. Korol, O. Rubchak, O. Orda, and D. Novytska, “Effectiveness of translation transformations in different styles of the English language for teaching written translation,” *Amazonia Investiga*, vol. 12, no. 67, pp. 185–197, 2023. <https://doi.org/10.34069/AI/2023.67.07.17>
- [7] Z. Feng and X. Wang, “Visual translation of English classroom teaching system based on mobile learning,” *International Journal of Continuing Engineering Education and Life Long Learning*, vol. 35, no. 8, pp. 58–73, 2025. <https://doi.org/10.1504/IJCEELL.2025.149039>
- [8] J. R. Fu and L. M. Cao, “A study of English translation teaching mode integrating China traditional culture,” *Voprosy Istorii*, vol. 5, no. 1, pp. 146–153, 2023. <https://doi.org/10.31166/VoprosyIstorii202305Statyi22>
- [9] M. Alieva, Y. Godis, I. Lazurenko, H. Konopelkina, and O. Lytvynko, “The use of translation transformations in different styles of the English language for teaching written translation,” *Apuntes Universitarios*, vol. 13, no. 3, pp. 92–104, 2023. <https://doi.org/10.17162/au.v13i3.1526>
- [10] S. Zubanova, E. Didenko, and I. Karabulatova, “Location-based mobile learning system facilitating English learning,” *Interactive Learning Environments*, vol. 31, no. 8, pp. 4818–4834, 2023. <https://doi.org/10.1080/10494820.2021.1983609>
- [11] I. Zulaeha, I. Subyantoro, C. Hasanudin, and R. Pristiwati, “Developing teaching materials of academic writing using mobile learning,” *Ingénierie des Systèmes d’Information*, vol. 28, no. 2, pp. 409–418, 2023. <https://doi.org/10.18280/isi.280216>
- [12] N. Terentieva, V. Karpenko, N. Yarova, N. Shkvyria, and M. Pasko, “Technological innovation in digital brand management: Leveraging artificial intelligence and immersive experiences,” *Journal of Research, Innovation and Technologies*, vol. 4, no. 2, pp. 201–223, 2025. [https://doi.org/10.57017/jorit.v4.2\(8\).06](https://doi.org/10.57017/jorit.v4.2(8).06)
- [13] X. Lyu, S. S. Ramasamy, and F. Ying, “Digital virtual anchors impact in entertainment industry: An exploration of user acceptance and market insights,” *Journal of Research, Innovation and Technologies*, vol. 4, no. 2, pp. 125–141, 2025. [https://doi.org/10.57017/jorit.v4.2\(8\).01](https://doi.org/10.57017/jorit.v4.2(8).01)

- [14] B. Nath, S. Sarkar, and S. Mukhopadhyay, “AstraMT: Instruction-tuned few-shot assamese–English translation with context-aware prompting and reranking,” *Computación y Sistemas*, vol. 29, no. 3, pp. 1167–1178, 2025. <https://doi.org/10.13053/cys-29-3-5886>
- [15] K. C. Hsu, N. E. Barrett, and G. Z. Liu, “English for tourism and AR-assisted context-aware ubiquitous learning: A preliminary design-based research study,” *Computer Assisted Language Learning*, vol. 38, no. 3, pp. 544–568, 2025. <https://doi.org/10.1080/09588221.2023.2202701>
- [16] E. Fabregas, G. Farias, S. Dormido-Canto, M. Guinaldo, J. Sánchez, and S. Dormido Bencomo, “Platform for teaching mobile robotics,” *Journal of Intelligent & Robotic Systems*, vol. 81, no. 1, pp. 131–143, 2016. <https://doi.org/10.1007/s10846-015-0229-8>
- [17] M. Gluchmanova, “Using the Moodle platform in English teaching,” *TEM Journal*, vol. 5, no. 4, pp. 492–497, 2016. <https://doi.org/10.18421/TEM54-13>

## 7 AUTHORS

**Xuejiao Liu** holds a Master’s degree. She is currently a Lecturer at Department of Basic Studies, Hebei University of Environmental Engineering, Qinhuangdao 066000, China, with research focus on foreign language education. With 13 years of teaching experience, she has presided over one provincial/ministerial-level research project and participated in several research projects funded by the Ministry of Education and other provincial-level authorities. She has published multiple papers on teaching reform (E-mail: [vanillalalaa@163.com](mailto:vanillalalaa@163.com)).

**Yinghua Hu** earned her bachelor’s degree from School of Foreign Languages, Yanshan University, Qinhuangdao 066004, China in 2007, followed by a master’s degree from Beijing Foreign Studies University in 2010. Since completing her graduate studies, she has served on the faculty at Yanshan University. Her scholarly contributions include two monographs and twelve research papers. Her primary research interest center on translation pedagogy and German language instruction (E-mail: [hu124yinghua@163.com](mailto:hu124yinghua@163.com)).

**Zhiyun Zhang** completed his undergraduate studies at Xi’an International Studies University, where he received his bachelor’s degree in 2008. From 2008 to 2011, he served on the faculty at West Yunnan University. From 2011 to 2014, he pursued a master’s degree at Shanghai International Studies University, graduating in 2010. Since then, he has been working at School of Foreign Languages, Yanshan University, Qinhuangdao 066004, China. He has published two monographs and sixteen research papers. His research interests primarily focus on translation teaching and translation studies (E-mail: [oliver0401011623@163.com](mailto:oliver0401011623@163.com)).