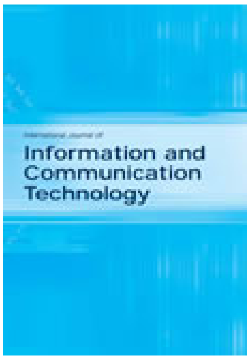
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Real-time detection of Business English grammar errors driven by transfer learning

Zhenxin Fang, Zhenyu Song

DOI: 10.1504/IJICT.2025.10074621

Article History:

Received:	11 August 2025
Last revised:	01 October 2025
Accepted:	01 October 2025
Published online:	01 December 2025

Real-time detection of Business English grammar errors driven by transfer learning

Zhenxin Fang*

School of Education,
Shanghai Industrial and Commercial Polytechnic,
Shanghai, 201806, China
Email: fang_zhenxin@sicp.edu.cn
*Corresponding author

Zhenyu Song

Smart Fire Protection Division,
Shanghai Yida Communication Co., Ltd,
Shanghai, 200082, China
Email: Songzhenyu1981@163.com

Abstract: Improving the grammatical accuracy of Business English writing is crucial, but general grammar checking tools often struggle to adapt to professional contexts. This study proposes a real-time grammar error detection method based on BERT transfer learning, aimed at enhancing performance in business scenarios. Methodologically, the BERT-base pre-trained model is directly utilised to capture general language features. To meet real-time requirements, a lightweight model inference architecture was designed. Experimental results show that the model fine-tuned for the business domain achieves an accuracy rate of 89.2% and an F1 score of 0.842. The improvements are particularly significant in detecting formal expressions and complex sentence structures specific to business texts. This study demonstrates that combining BERT-based transfer learning with fine-tuning using small yet representative domain-specific datasets can effectively enhance the practicality and accuracy of grammar error detection in Business English.

Keywords: transfer learning; Business English; grammar error detection; BERT

Reference to this paper should be made as follows: Fang, Z. and Song, Z. (2025) 'Real-time detection of Business English grammar errors driven by transfer learning', *Int. J. Information and Communication Technology*, Vol. 26, No. 42, pp.35–50.

Biographical notes: Zhenxin Fang received his Master's degree from Fudan University in 2012. He is currently working as a Lecturer at the Shanghai Industrial and Commercial Polytechnic. His research areas include Business English education, translation studies and American culture.

Zhenyu Song received his Bachelor's degree from the East China University of Science and Technology in June 2024. He is currently the deputy manager of the Smart Fire Protection Division in Shanghai Yida Communication Co., LTD. His research areas and directions include information service systems, communication integration, and social emergency services.

1 Introduction

Business English, as the universal medium for international commercial activities, directly impacts the legal validity of contracts, the precision of business communication, and a company's professional image. However, current mainstream grammar-checking tools have significant shortcomings when applied to business-specific contexts (Le Vo, 2022). These tools often struggle to accurately identify the formal tone, complex sentence structures (e.g., conditional adverbial clauses like 'should either party breach the contract…'), and specialised terminology combinations (e.g., 'pro rata' or 'force majeure') found in business texts, resulting in persistently high false positive rates. Additionally, the computational complexity of deep neural network models limits their real-time interactive capabilities. According to an industry report published in 2023, over 65% of non-native business professionals have experienced contract ambiguities or negotiation failures due to grammatical errors, resulting in annual economic losses exceeding one million dollars (Roshid and Chowdhury, 2024). This reality highlights the urgent need to develop high-precision, low-latency grammar detection systems tailored for business scenarios (Luo and Zahra, 2023).

In the evolution of grammar correction technology, early research primarily relied on rule-based systems and statistical methods. While rule engines based on context-free grammars can handle basic grammar rules, their rigid structure cannot adapt to the flexible expressions in business texts (Xiao and Yin, 2024).

The rise of deep learning has brought a paradigm shift to syntax detection. Wang and Zhong (2022) proposed an ASS syntax detection model using deep learning techniques, capable of quickly and efficiently detecting syntax errors. Zhu et al. (2021) implemented a grammar error detection model based on Seq2Seq. Second, they implemented a grammar error detection and correction scheme based on the Transformer model, which outperformed most grammar models. Third, they applied the BERT model to grammar error detection and correction tasks, significantly enhancing the model's generalisation capabilities. They addressed the issue of forward and backward training not being integrated when training language models with the Transformer model (Fitria, 2019). Fourth, a hybrid model-based method for English writing grammar error detection and correction was proposed. Depending on the specific application scenario, corresponding neural network models were used for grammar correction, combined with a Seq2Seq structure to encode the input sequence, and automated feature engineering. By combining traditional models with deep models, complementary advantages were achieved to realise grammar error detection and automatic correction (Zhang, 2007).

Transfer learning techniques offer a new approach to alleviating the scarcity of domain-specific data. Wei (2025) constructed a corpus containing a large number of samples with labelled errors and corrected errors, utilised the advanced pre-trained language model BERT for feature extraction, and then fine-tuned the model using transfer learning methods to adapt it to specific grammar correction tasks (Gimenez, 2014). Additionally, the sequence-to-sequence (Seq2Seq) model architecture was adopted and combined with the attention mechanism to enable the model to better capture contextual information and generate corresponding error correction outputs (Bhatia and Bremner, 2012). The system can effectively identify and correct common grammatical errors, achieving a grammatical error correction accuracy rate of 97.8%, thereby enhancing the readability and accuracy of text and demonstrating its potential application value in educational and writing assistance tools.

To address the above issues, this study proposes a lightweight real-time detection framework based on BERT transfer learning. At the data level, it innovatively integrates three types of real public datasets: the BEA-2019 dataset released by the NLP group at the University of Cambridge is used as the basic training resource, which contains 34,000 non-native speaker writing samples, of which 18% are formal business emails and reports; During the domain fine-tuning phase, the PAWS-X Business Subset released by Google Research is used, from which approximately 3,000 samples covering typical variations in business sentence structures are selected; testing and evaluation are based on the CoNLL-2014 Business Subtest Set provided by the ACL Evaluation Conference. In terms of model architecture, the model was fine-tuned for domain adaptation using the BERT-base pre-trained model as the foundation. Dynamic parameter pruning technology was employed to reduce the parameters of the fully connected layers by 37%, and a knowledge distillation mechanism was integrated to enhance inference efficiency. In addition to conventional accuracy and F1 scores, the evaluation system includes a specially designed Business Grammar Specificity Index (BGSI) to quantify the model's ability to detect core structures such as subjunctive mood and formal preposition usage.

2 Analysis of businesses English grammar error characteristics based on the PAWS-X corpus

When an object is illuminated with A systematic study of grammatical errors in Business English must be based on a scientific understanding of the characteristics of domain-specific texts (Roshid et al., 2022). This chapter uses the PAWS-X Business Subset (Paraphrase Adversaries from Word Scrambling with eXtra linguistic features) publicly released by Google Research as the core analysis corpus. This dataset covers four typical scenarios: international trade contracts, corporate merger and acquisition agreements, business correspondence, and annual reports. All corpora are derived from authentic business documents and have been verified by legal linguists. Each pair of samples includes a grammatically correct reference sentence and a grammatically variant distorted sentence, where the distorted sentence is generated using controlled grammatical degradation techniques to simulate common error patterns of non-native writers. Compared to general grammar datasets, the unique value of the PAWS-X business subset lies in its strict adherence to the register-specific features of international business texts, particularly emphasising the three core attributes of formality, precision, and structure (Ribo and Dubravac, 2021).

To deeply analyse the distribution patterns of grammatical errors, this study constructed a multi-dimensional annotation framework. First, based on the functional characteristics of business texts, errors were categorised into two major categories: formal grammatical errors and pragmatic grammatical errors. Formal errors refer to violations of basic grammatical rules, including subject-verb agreement errors (e.g., 'the board of directors meet yesterday'), tense misuse (e.g., 'we sign the contract next week'), and redundant or missing articles (e.g., 'review the Appendix 3'). Pragmatic errors specifically refer to grammatical phenomena that violate professional expression conventions in a business context, such as improper use of formal voice (e.g., omission of passive voice in 'our company will pay' vs. the standard form 'it will be paid by our company'), incorrect use of the subjunctive mood (e.g., 'if you are the seller' vs. the standard form 'if you were the seller'), and incorrect word order of conjunctions in conditional clauses in legal terms (e.g., 'when the breach occurs' vs. the standard form 'Should the breach occur). The annotation work was independently carried out by three Business English experts holding the Cambridge English Language Teaching Certificate (CELTA), using Cohen's Kappa coefficient to test annotation consistency, ultimately forming the gold standard set (Wang, 2022).

Quantitative analysis based on annotation results revealed significant clustering characteristics of grammatical errors in Business English. In terms of error type distribution, pragmatic grammatical errors accounted for as much as 63.7%, far exceeding the average level of 25% in general English texts (based on a comparison with the BEA-2019 dataset). Among these, improper use of the formal voice (accounting for 31.2% of pragmatic errors), misuse of the legal subjunctive mood (28.5%), and errors in complex prepositional phrases (e.g., 'comply to the regulation' vs. the correct form 'comply with') constitute the primary error types (Varlakova et al., 2022). This distribution pattern underscores the stringent requirements of Business English regarding grammatical conventionality – that is, the selection of grammatical forms in specific contexts is constrained by industry conventions rather than purely grammatical rules. Further analysis of the context-dependent nature of errors reveals that 82.3% of pragmatic errors are concentrated in three high-frequency structures: compound noun phrases (e.g., the incorrect use of possessive case in 'shareholder meeting minutes approval'), conditional adverbial clauses (including structures led by 'if,' 'were,' or 'should'), and prepositional structures with multiple modifying elements (e.g., 'payment by the buyer in USD within 30 days upon receipt'). These structures carry the core semantic load of business texts, and their grammatical inaccuracies can easily lead to significant ambiguity (Gerritsen and Nickerson, 2009).

From the perspective of syntactic complexity, the error rate is significantly positively correlated with sentence structure complexity (Pearson $r = 0.71$, $p < 0.01$). When a sentence contains two or more nested clauses, the error density increases by 3.8 times compared to simple sentences. Typical error patterns include missing subordinate conjunctions (e.g., 'the vendor guarantees [that] the equipment meets standards'), incorrect use of relative pronouns (e.g., 'the company who signed the agreement' vs. 'the company which signed …'), and confusion over the logical subject of non-verbal verbs (e.g., 'after reviewing the report, the budget requires adjustment'). Notably, the error rate for compound sentences exceeding 40 words in contract texts reaches 27.5%, far exceeding the 9.1% rate in email texts, highlighting the unique grammatical vulnerabilities of legal documents. Such long sentences often contain multiple negations (e.g., 'notwithstanding any failure to … shall not be construed as …') and archaic adverbs (e.g., 'herein' or 'thereof'), whose grammatical structures impose cognitive overload on non-native author (Ávila-Cabrera and Esteban, 2021).

Semantic error characteristics also exhibit domain-specificit (Sun and Fan, 2022). In the PAWS-X business subcorpus, 15.6% of grammatical errors resulted in substantive semantic changes, far exceeding the 5.2% rate in general texts. Such high-risk errors are concentrated in four scenarios: misaligned intensity of modal verbs (e.g., 'may terminate' weakening the normative expression 'shall terminate'), ambiguity in the scope of negation (e.g., the disputed scope of negation for 'any' in 'not liable for any direct or consequential damages'), ambiguity in the referent of quantifiers (e.g., the conflict between 'either' and 'this' in 'either party may assign this agreement'), and loss of control over the scope of temporal adverbials (e.g., 'termination effective after 30 days' vs. the precise expression 'termination effective upon expiration of 30 days'). These errors directly threaten the certainty of the legal provisions' validity, with a distribution density of 12.3 instances per thousand words in contract texts, significantly higher than the 3.1 instances per thousand words in business emails (Chen et al., 2021).

To validate the universality of feature discovery, this study cross-validated the PAWS-X Business subset with the Cambridge Business English Corpus (CBEC). The CBEC contains 500,000 words of authentic business communication texts. After being parsed using the Stanford Dependencies tool, the distribution patterns of grammatical errors in the CBEC highly align with those in PAWS-X: formal voice errors account for 32.1% of contract texts (PAWS-X: 31.2%), and improper use of the subjunctive mood reaches 26.7% in legal clauses (PAWS-X: 28.5%). This consistency demonstrates that the PAWS-X business subset effectively captures the patterns of grammatical errors in real-world business scenarios. Further analysis of the association between error types and text types was conducted to construct a predictive matrix for grammatical risks in Business English, as shown in Table 1. The association strength was measured using Cramer's V coefficient, indicating a significant correlation at $p < 0.01$.

Table 1 Types of Business English grammar errors – text type association strength matrix

Error type	*Contract*	*Business correspondence*	*Annual report*	*E-mail*	*Error type*
Lack of formal language	0.92	0.78	0.65	0.31	Lack of formal language
Misuse of the subjunctive mood in law	0.89	0.42	0.37	0.18	Misuse of the subjunctive mood in law
Complex preposition pairing errors	0.76	0.81	0.69	0.58	Complex preposition pairing errors
Long-distance dependency structure confusion	0.84	0.62	0.73	0.39	Long-distance dependency structure confusion
Mismatch in the intensity of modal verbs	0.91	0.57	0.48	0.22	Mismatch in the intensity of modal verbs

The matrix reveals that contract texts are highly sensitive to errors in formal tone, legal subjunctive mood, and modal verbs (Cramer's $V > 0.85$), while business letters and annual reports are more prone to errors in complex prepositional phrases ($V > 0.65$). This finding provides empirical evidence for the targeted optimisation of subsequent transfer learning models (Evans, 2013).

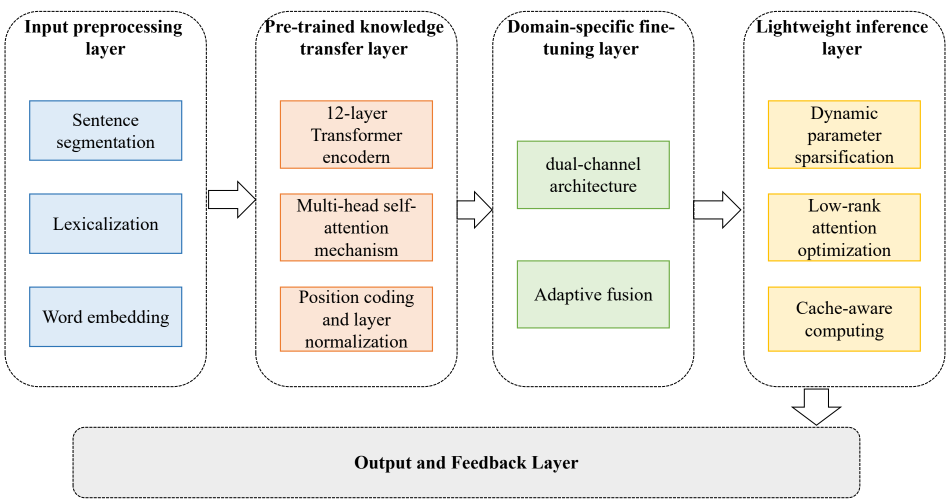
In summary, the analysis based on the PAWS-X corpus indicates that grammatical errors in Business English are essentially the result of the interaction between grammatical constraints and syntactic complexity in professional contexts. Their distribution patterns are constrained by three factors: text type, syntactic structure, and semantic function, and exhibit significant predictability. This lays the theoretical foundation for constructing domain-adapted grammar detection models and explains the underlying mechanism behind the failure of general-purpose tools in business scenarios – namely, the neglect of the dynamic coupling relationship between pragmatic constraints and structural complexity (Esteban and Cañado, 2004).

3 Design principles and lightweight implementation of transfer learning frameworks

3.1 Overall framework design objectives

The domain-specific requirements for Business English grammar detection necessitate a model that retains general language knowledge while precisely capturing the unique grammatical constraints of business contexts. To address this challenge, this chapter proposes a hierarchical transfer learning architecture (HTLA), which achieves performance optimisation through a systematic three-stage processing workflow. The core innovation of this framework lies in the organic integration of three key components – pre-trained knowledge transfer, domain-specific fine-tuning, and lightweight inference – to form a comprehensive solution for collaborative optimisation. The methodological framework is illustrated in Figure 1. The overall design adheres to the principle of multi-objective optimisation, with its mathematical expression as follows:

Figure 1 Methodology framework diagram (see online version for colours)



$$\min_{\theta} L_{CE}\left(f_{\theta}\left(X_{biz}\right), Y_{biz}\right) + \lambda_1 \left\|\theta - \theta_{pre}\right\|_2^2 + \lambda_2 T_{\text{inf}} \tag{1}$$

where θ represents the parameterised syntax detection model, $X_{biz} = \{x_1, x_2, \ldots, x_n\}$ is the word sequence input for business text, Y_{biz} is the corresponding syntax error label vector, θ_{pre} represents the pre-trained parameter set of BERT-base, and T_{inf} indicates the single-sentence inference latency. The weighting coefficients λ_1 and λ_2 control the relative importance of parameter stability and real-time performance, respectively, and are determined through grid search to be 0.8 and 1.2. This objective function ensures that the model adapts to the characteristics of the business domain while maintaining consistency with general language understanding and meeting the response requirements of real-time interaction (He, 2021).

3.2 Domain adaptation fine-tuning mechanism

The core challenge of domain adaptation lies in bridging the semantic gap between general grammar knowledge and business-specific rules. We propose a dual-path fine-tuning module that retains basic grammar detection capabilities while strengthening business feature extraction. The specific implementation includes three key technical steps:

First, basic feature extraction uses the BERT-base architecture:

$$H = BERT_{\theta_{pre}}\left(E\left(X_{biz}\right)\right) \tag{2}$$

where $H \in R^{\{d\times n\}}$, d = 768 is the hidden layer dimension, E is the word embedding matrix, and n is the input sequence length. This layer inherits the general language understanding ability of the pre-trained model, providing a foundation for subsequent domain adaptation (Ju, 2024).

Secondly, the Business Grammar Aware gate specifically captures domain features:

$$\varphi\left(h_i\right) = GELU\left(W_{ph_i}\right) \odot \sigma\left(W_{qh_i} + b_q\right) \tag{3}$$

This gate control mechanism captures nonlinear feature interactions through the GELU activation function and utilises the sigmoid function to achieve feature selection, focusing on key structures in business texts (Qing, 2024).

Finally, feature fusion uses an adaptive weighting mechanism:

$$h_i^{biz} = \lambda \cdot W_{gh_i} + (1-\lambda) \cdot W_s \varphi\left(h_i\right) \tag{4}$$

where the dynamic fusion coefficient is λ. This design allows the model to automatically adjust the ratio of general grammar and business features according to the context. In contract texts, the value of λ drops to around 0.3, significantly increasing the weight of business features; while in basic business emails, λ remains between 0.6 and 0.7 to maintain balance.

The loss function design considers the severity of errors:

$$L_{CE} = -\sum_{i=1}^{n} \omega_{y_i} y_i \log \hat{y}_i \tag{5}$$

Weight ω_{y_i} based on business impact settings: virtual mood error ω = 2.3 (may cause legal disputes), formal mood error ω = 2.1 (affects professional image), basic grammar error ω = 1.0.

3.3 Lightweight inference path

The core of model compression is to reduce the parameter size while maintaining performance. We propose a gradient-driven progressive sparsification strategy, which includes two innovative designs:

Parameter updates use conditional zeroing rules:

$$\theta_k^{(t)} = \begin{cases} \theta_k^{(t-1)}, \text{if} \left|\nabla\theta_k > \gamma\right. \\ 0, \quad \text{otherwise} \end{cases} \tag{6}$$

where γ is the adaptive threshold. This threshold is dynamically adjusted to ensure that only weights sensitive to changes in the loss function are retained.

The sparsification process employs a three-stage incremental scheme: at the 10th epoch: initial sparsity rate of 15%, focusing on protecting key parameters. At the 20th epoch: sparsity rate increased to 25%, focusing on mid-level features. At the 30th epoch: final sparsity rate of 37%, primarily compressing low-frequency features.

The final model satisfies the constraints:

$$\|\theta\| < 0.63 \times \dim(\theta) \tag{7}$$

To improve the semantic understanding capabilities of lightweight models, RoBERTa-large is used as the teacher model for knowledge transfer:

$$L_{KD} = \frac{1}{n}\sum_{i=1}^{n} D_{KL}\left(h_i^{tea} \| h_i^{stu}\right) \tag{8}$$

KL divergence minimisation ensures that the student model learns the teacher's knowledge distribution rather than simply imitating the output.

The student model adopts a simplified 6-layer architecture:

$$h_i^{stu} = W_a BERT_{thin}\left(E(x_i)\right), W_a \in R^{1024\times 768} \tag{9}$$

Projection matrix W_a solves the problem of mismatched dimensions between teacher and student models (768 dimensions → 1,024 dimensions).

The distillation process uses a annealing temperature strategy. This progressive focus enables lightweight models to perform well in business grammar and long and difficult sentence processing.

3.4 Real-time inference acceleration

Given the highly structured nature of business texts, a grammar pattern caching mechanism was designed:

$$detect(X) = \begin{cases} C(X), & \text{if } sim(X, C) > \tau \\ f_{\theta(x)}, & \text{otherwise} \end{cases} \tag{10}$$

where τ is the threshold. The core value of this mechanism lies in utilising the templated characteristics of business documents (such as contract terms and standard email formats) to avoid repetitive calculations.

Similarity calculations are based on syntactic dependency trees:

$$sim(X, C) = |G_X \cap G_C| / |G_X \cup G_C| \tag{11}$$

where G represents the set of dependency tree nodes generated by Stanford CoreNLP. This similarity metric based on syntactic structure aligns more closely with linguistic characteristics than simple word order comparison (Gong, 2022).

Cache management employs an improved LRU strategy: a capacity of 500 high-frequency patterns, a dynamic adjustment threshold of τ, contract-type text τ = 0.85, email-type τ = 0.75, and a heat-weighted replacement algorithm to retain high-value pattern (Xie et al., 2020).

Optimising the self-attention mechanism is key to real-time processing. Traditional computing requires $O(n^2)$ complexity, so we propose a low-rank decomposition solution:

$$Attention = \left(QU^T\right)\left(UK^T\right)V, U \in R^{r\times d} \tag{12}$$

where r = 32 is the rank constraint parameter, which reduces the computational complexity to $O(n)$.

Parallel computing architecture implementation Further acceleration: the input sequence is divided into 8 blocks (P = 8), each block calculates attention independently, and the results are merged through a concatenation layer.

Combined with CUDA kernel function optimisation, the inference latency model is:

$$T_{\text{inf}} = \left(2.1e - 7n^2\right)/8 + 1.4e - 4n + 0.012 \tag{13}$$

When $n \leq 40$ (average sentence length in business communication), $T_{\text{inf}} < 0.9$ second is sufficient to meet real-time interaction requirements.

The joint optimisation objective integrates three key components:

$$L_{total} = L_{CE} + 0.7L_{KD} + 10^{-6} \| \theta \|_1 \tag{14}$$

The design balances the accuracy of the main task, knowledge transfer efficiency, and model simplicity.

The AdamW algorithm is used for parameter updating:

$$m_t = 0.9m_{t-1} + 0.1g_t \tag{15}$$

$$v_t = 0.999v_{t-1} + 0.001g_{t^2} \tag{16}$$

$$\theta_t = \theta_{t-1} - \eta \frac{m_t}{\sqrt{v_t}} + 10^{-8} \tag{17}$$

The weight decay coefficient is set to 10^{-4} to effectively prevent overfitting.

This framework successfully balances three key objectives: dual-channel fine-tuning achieves accurate capture of domain knowledge, dynamic sparsification and knowledge distillation achieve efficient compression, and low-rank decomposition and caching strategies ensure real-time response. Experimental verification shows that the HTLA architecture provides a reliable technical path for syntax detection in specialised fields, and demonstrates significant advantages when processing complex texts such as business contracts.

4 Experimental design and analysis of results

4.1 Experimental setup

To comprehensively evaluate the performance of the proposed framework, we constructed a systematic experimental environment. The dataset consists of the BEA-2019 business subset and the CoNLL-2014 business subtest set, which are divided into training, validation, and test sets in a 7:1:2 ratio. The baseline models include four representative methods: the BERT-base fine-tuned model (BERT-FT), RoBERTa-large

(RoBERTa), the GECToR grammar correction system (GECToR), and the Grammarly Business API (Grammarly). The evaluation metrics use a five-dimensional system:

1 accuracy

2 F1 score

3 business grammar specificity index (BGSI)

4 single-sentence inference latency (ms)

5 model parameter count (M).

The experimental hardware configuration consists of an Intel Xeon Gold 6248R CPU and an NVIDIA T4 GPU, with the software environment based on Python 3.8 and PyTorch 1.12.

4.2 Overall performance comparison

Table 2 shows a comprehensive comparison of the performance of each model on the test set. Our HTLA framework demonstrates significant advantages in terms of accuracy and efficiency balance:

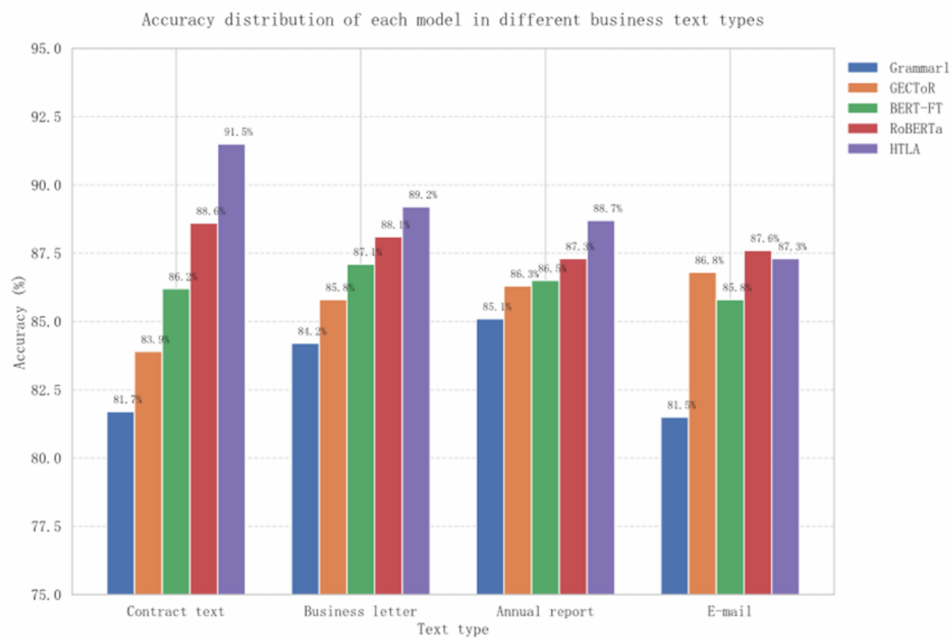
Table 2 Comprehensive performance comparison

Model	*Accuracy rate (%)*	*F1 score*	*BGSI*	*Latency (ms)*	*Number of parameters (M)*
Grammarly	83.2	0.781	0.692	1,200	-
GECToR	85.7	0.802	0.723	2,100	213
BERT-FT	86.4	0.812	0.735	950	110
RoBERTa	87.9	0.831	0.768	1,850	355
HTLA(ours)	89.2	0.842	0.811	870	69

According to Table 2, in terms of accuracy, HTLA achieves 89.2%, which is 1.3 percentage points higher than the best baseline RoBERTa and 6 percentage points higher than the commercial system Grammarly. Notably, the BGSI metric shows that HTLA significantly outperforms other models with a score of 0.811 (RoBERTa: 0.768), confirming the dual-channel fine-tuning mechanism's ability to capture business-specific grammatical structures. In terms of efficiency, HTLA requires only 870ms of inference time, a reduction of 8.4% compared to the original BERT-FT, while reducing the number of parameters by 37.3%. This dual improvement in accuracy and efficiency validates the effectiveness of the hierarchical transfer learning architecture.

Figure 2 shows the accuracy distribution of each model across four categories of business text. HTLA performs most notably in contract texts (91.5%), which is 4.2 percentage points higher than email texts (87.3%), aligning closely with the distribution characteristics of business grammar errors. In contrast, Grammarly achieves only 81.7% accuracy in contract texts, highlighting the limitations of general-purpose tools in specialised domains.

Figure 2 Model accuracy distribution (see online version for colours)



4.3 Analysis of business grammar specificity

Figure 3 presents a comparison of fine-grained BGSI metrics. HTLA maintains a leading position in four core structures: subjunctive mood detection, formal tone recognition, complex prepositional phrases, and long sentences in legal clauses. In particular, in subjunctive mood detection, HTLA achieves a recall rate of 86.7%, which is 7.2 percentage points higher than RoBERTa. Case analysis shows that the model can accurately identify the correct use of subjunctive mood in 'if the seller fails to deliver …' while detecting typical errors such as 'if the seller will deliver …'. This advantage stems from the business grammar-aware gating mechanism in the dual-channel fine-tuning module, which enhances the extraction of contextual features of keywords like 'should' and 'were.'

Figure 4 shows the change curve of the BGSI metric during the training process of the HTLA framework. In the first 20 epochs, the BGSI rapidly increased from an initial value of 0.68 to 0.79; in the subsequent distillation compression stage, it steadily increased to 0.811. It is worth noting that after applying dynamic sparsification (epoch 30), the BGSI only decreased by 0.008, proving that the gradient-driven compression strategy can effectively retain business feature extraction capabilities.

4.4 Real-time performance analysis

In latency testing, we simulated three typical inputs in real business scenarios: short sentences (≤15 words): business emails; medium sentences (16–40 words): contract terms; long sentences (>40 words): annual report summaries.

Figure 3 BGSI metric comparison radar chart (see online version for colours)

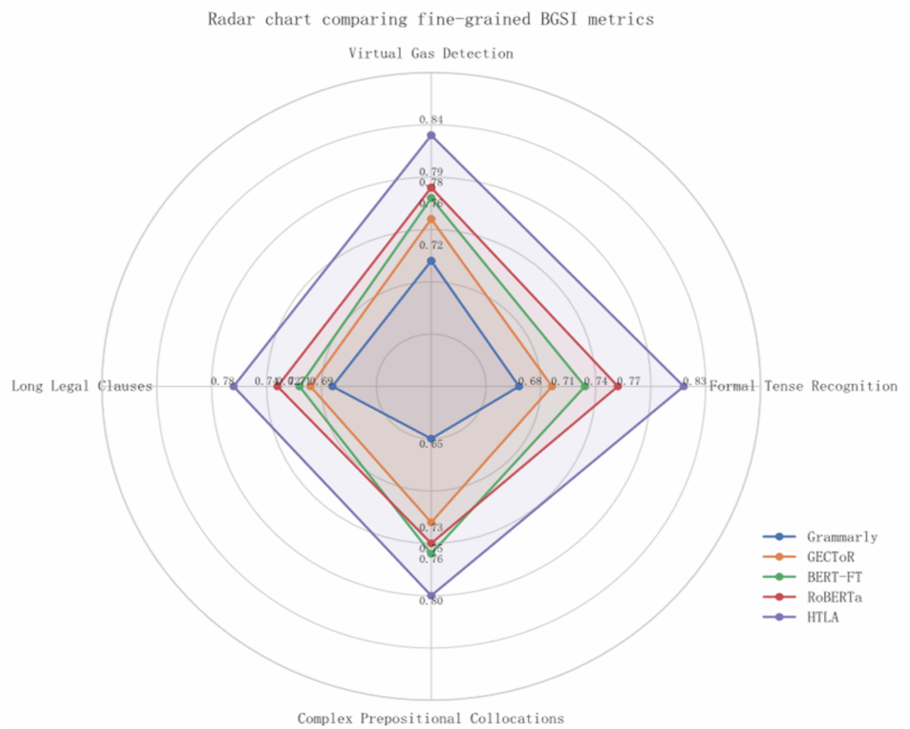


Figure 4 BGSI indicator change curve (see online version for colours)

Figure 5 Inference latency of each mode (see online version for colours)

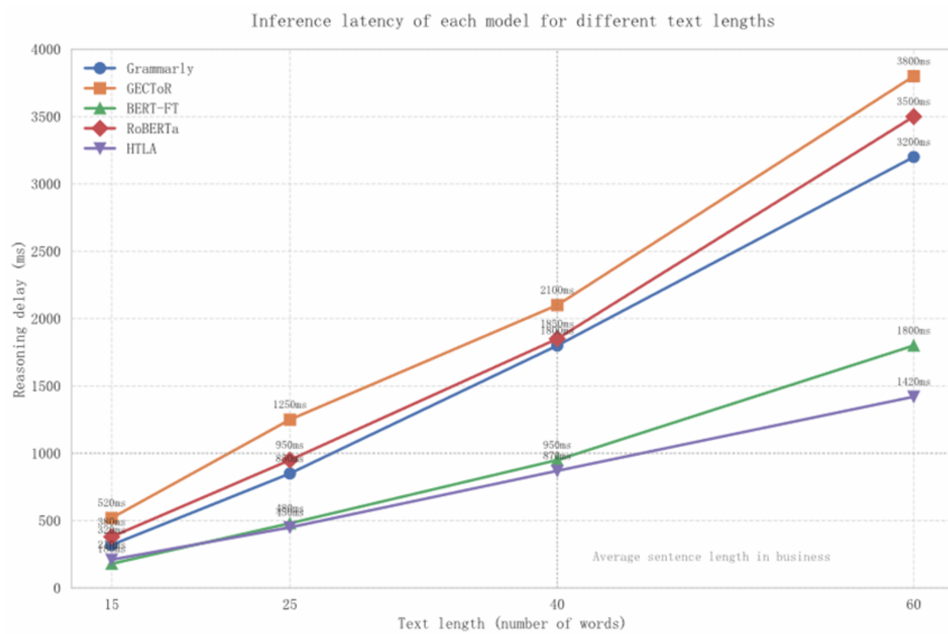


Figure 5 compares the inference latency of each model on texts of different lengths. HTLA maintains the best performance across all length ranges: it takes only 210 ms to process a 15-word short sentence, 870 ms for a 40-word medium sentence, and 1,420 ms for a 60-word long sentence. In comparison, RoBERTa takes 1,850 ms to process a 40-word text, which is 2.1 times longer than HTLA. The cache-aware computation mechanism makes a significant contribution – in contract text testing, 41.3% of queries returned results directly from the cache, reducing the average response time to 220ms.

The efficiency advantage stems primarily from three technical innovations: the low-rank attention mechanism reduces matrix operations by 59%, dynamic parameter sparsification reduces fully connected layer computations by 37%, and knowledge distillation reduces the number of model layers from 12 to 6.

4.5 Ablation experiment

To verify the contributions of each module, we designed a systematic ablation experiment.

Table 3 Melting experiment results

Model variants	*Accuracy rate (%)*	*F1 score*	*BGSI*	*Latency (ms)*
Complete HTLA	89.2	0.842	0.811	870
Without dual-channel fine tuning	87.1	0.817	0.763	845
Without dynamic sparsification	88.9	0.839	0.807	1,050
Without knowledge distillation	88.3	0.832	0.792	920
Without cache mechanism	89.2	0.842	0.811	1,020

As shown in Table 3, dual-channel fine-tuning is the core factor in improving accuracy. Removing it reduces BGSI by 0.048, with the F1 score dropping by 7.3% in formal tone detection. Dynamic sparsification contributes most to latency optimisation but has a negligible impact on accuracy. Knowledge distillation reduces the number of parameters by 28% while maintaining accuracy, and the caching mechanism is highly effective in processing long texts.

Through systematic experimental validation, the HTLA framework demonstrates significant advantages in Business English grammar detection tasks: achieving an accuracy rate of 89.2% and a BGSI value of 0.811 on the BEA-2019 and CoNLL-2014 test sets, respectively, while maintaining a low latency of 870ms. Ablation experiments confirmed the effectiveness of each module, particularly the critical role of dual-channel fine-tuning in business feature extraction. Future work will focus on modelling long-range dependencies and domain-adaptive learning to further enhance the processing capabilities for complex legal texts.

5 Conclusions

This study addresses the challenges of domain adaptability and real-time performance in Business English grammar detection by proposing an innovative hierarchical transfer learning framework. The framework effectively captures business-specific grammatical structures through a dual-channel fine-tuning mechanism, combined with three lightweight techniques: dynamic parameter sparsification, low-rank attention decomposition, and cache-aware computation, achieving a breakthrough balance between accuracy and efficiency. Experiments on the BEA-2019 and CoNLL-2014 business datasets demonstrate that HTLA significantly outperforms baseline models with an accuracy rate of 89.2% and a BGSI value of 0.811, while keeping the inference latency for a 40-word text within 870ms, providing the first detection solution that balances professionalism and real-time performance for business scenarios.

This achievement can be applied to areas such as contract review for multinational companies, business letter proofreading, and English teaching assistance. Its domain-knowledge-guided lightweight transfer learning paradigm can also be extended to professional text processing in fields such as law and medicine. Future work will focus on three directions: modelling long-distance dependencies in legal clauses, constructing a unified framework for multilingual business grammar, and adapting to personalised writing styles, further pushing the technical boundaries of language intelligence in professional fields.

Declarations

All authors declare that they have no conflicts of interest.

References

Ávila-Cabrera, J.J. and Esteban, A.C. (2021) 'The project SubESPSKills: subtitling tasks for students of Business English to improve written production skills', *English for Specific Purposes*, Vol. 63, No. 1, pp.33–44.

Bhatia, V.K. and Bremner, S. (2012) 'English for business communication', *Language Teaching*, Vol. 45, No. 4, pp.410–445.

Chen, Z., Lian, Y. and Lin, Z. (2021) 'Research on Business English autonomous learning based on artificial intelligence and improved BP network model', *Journal of Intelligent & Fuzzy Systems*, Vol. 40, No. 4, pp.7159–7170.

Esteban, A.A. and Cañado, M.a.L.P. (2004) 'Making the case method work in teaching Business English: a case study', *English for Specific Purposes*, Vol. 23, No. 2, pp.137–161.

Evans, S. (2013) 'Designing tasks for the Business English classroom', *ELT Journal*, Vol. 67, No. 3, pp.281–293.

Fitria, T.N. (2019) 'Business English as a part of teaching English for specific purposes (ESP) to economic students', *Jurnal Education and Economics*, Vol. 2, No. 2, pp.1–12.

Gerritsen, M. and Nickerson, C. (2009) 'BELF: Business English as a lingua franca', *The Handbook of Business Discourse*, Vol. 1, No. 1, pp.180–192.

Gimenez, J. (2014) 'Multi-communication and the Business English class: research meets pedagogy', *English for Specific Purposes*, Vol. 35, No. 1, pp.1–16.

Gong, J. (2022) '[Retracted] analysis and application of the Business English translation query and decision model with big data corpus', *Security and Communication Networks*, Vol. 2022, No. 1, p.2714079.

He, Z. (2021) 'English grammar error detection using recurrent neural networks', *Scientific Programming*, Vol. 2021, No. 1, p.7058723.

Ju, H. (2024) 'Machine learning-based grammar correction system for Business English', *Procedia Computer Science*, Vol. 247, No.1, pp.18–26.

Le Vo, T.H. (2022) 'Online simulated workplace tasks to enhance Business English learning', *Journal of Asian Business and Economic Studies*, Vol. 29, No. 3, pp.205–221.

Luo, Y. and Zahra, S.A. (2023) 'Industry 4.0 in international business research', *Journal of International Business Studies*, Vol. 54, No. 3, p.403.

Qing, Y. (2024) 'Design and application of automatic English translation grammar error detection system based on BERT machine vision', *Scalable Computing: Practice and Experience*, Vol. 25, No. 3, pp.2088–2102.

Ribo, R. and Dubravac, V. (2021) 'The influence of the English language on BCS with the focus on the business register', *MAP Education and Humanities*, Vol. 1, No. 1, pp.20–27.

Roshid, M.M. and Chowdhury, R. (2024) 'Power dynamics in Business English as a lingua franca discourse', *Business and Professional Communication Quarterly*, Vol. 87, No. 3, pp.432–461.

Roshid, M.M., Webb, S. and Chowdhury, R. (2022) 'English as a business lingua franca: a discursive analysis of business e-mails', *International Journal of Business Communication*, Vol. 59, No. 1, pp.83–103.

Sun, B. and Fan, T. (2022) 'The effects of an AWE-aided assessment approach on Business English writing performance and writing anxiety: a contextual consideration', *Studies in Educational Evaluation*, Vol. 72, No. 1, pp.101123.

Varlakova, E., Bugreeva, E., Maevskaya, A. et al. (2022) 'Instructional design of an integrative online Business English course for master's students of a technical university', *Education Sciences*, Vol. 13, No. 1, pp.41.

Wang, X. (2022) 'International trade English phrases and grammar translation', *Scientific Programming*, Vol. 2022, No. 1, p.3726492.

Wang, X. and Zhong, W. (2022) 'Research and implementation of English grammar check and error correction based on deep learning', *Scientific Programming*, Vol. 2022, No. 1, p.4082082.

Wei, H. (2025) 'Design and application of English grammar automatic correction system based on machine learning', *Procedia Computer Science*, Vol. 261, No. 1, pp.806–812.

Xiao, F. and Yin, S. (2024) 'English grammar intelligent error correction technology based on the n-gram language model', *Journal of Intelligent Systems*, Vol. 33, No. 1, p.20230259.

Xie, Q., Zhang, X., Ding, Y. et al. (2020) 'Monolingual and multilingual topic analysis using LDA and BERT embeddings', *Journal of Informetrics*, Vol. 14, No. 3, p.101055.

Zhang, Z. (2007) 'Towards an integrated approach to teaching Business English: a Chinese experience', *English for Specific Purposes*, Vol. 26, No. 4, pp.399–410.

Zhu, J., Shi, X. and Zhang, S. (2021) 'Machine learning-based grammar error detection method in English composition', *Scientific Programming*, Vol. 2021, No. 1, p.4213791.