DUAL-BRANCH CHINESE SPELLING CORRECTION METHOD FOR TARGET PERSONAL NAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Chinese personal name correction is an important problem in Chinese spelling correction (CSC) and has been widely applied in search correction, writing assistance, and post-processing of recognition results. Existing name correction methods often rely too heavily on name boundary detection, candidate name sets, or external knowledge resources, making them susceptible to biases caused by errors in obtaining candidate names and insufficient resource coverage. To address these problems, we incorporate Chinese personal name correction into a unified CSC framework and propose a Dual-Branch Chinese Spelling Correction (DCSC) method for the character-level correction of target personal names. Firstly, an online correction dataset for Chinese target person names is constructed. This dataset was created by replacing the personal pronouns in the existing Chinese corpus and generating a series of erroneous name samples, thereby avoiding the problem of incorrect candidate names due to errors in name boundary detection.Furthermore, we present a dual-branch gated fusion module. This module consists of a name-specific branch and a general branch, which respectively model the name-specific features and the general correction features. Through this module, Chinese personal name correction and general CSC are integrated into a single framework.In addition, a gating mechanism is employed to automatically fuse the features obtained by the two branches, thereby reducing the interference of the training for specific names on the overall correction ability of the model.DCSC not only can accurately correct errors in personal names, but also does not significantly reduce the overall spelling correction performance of the model. On the three classical CSC models, namely, MDCSpell, SCOPE, and NamBert, our method achieved F1 scores of 94.14%, 95.10%, and 93.95%, respectively, on the proposed online Chinese personal name correction dataset. This indicates that our method can effectively correct both Chinese personal names and common spelling errors simultaneously.

1 INTRODUCTION

As a fundamental task in Chinese natural language processing, Chinese Spelling Correction (CSC) aims to detect and correct erroneous characters in Chinese text. Current research (Wu et al., 2024) has explored methods such as error detection, candidate generation, phonological(Zhang et al., 2026) and visual feature modeling, and multimodal fusion(Li et al., 2025), achieving strong performance in general-domain correction. However, spelling errors also occur in low-frequency named entities such as personal names, locations, and organization names, where generic CSC models still face significant challenges(Liu et al., 2025a).

This paper focuses on spelling errors in Chinese personal names. Similar to errors in common vocabulary, name errors are typically caused by substitutions of homophonic, near-phonetic, or visually similar characters. However, personal names differ significantly from ordinary words in linguistic characteristics. Common words usually exhibit relatively stable semantics and collocational patterns, enabling models to assess their plausibility based on context. In contrast, personal names are sparsely distributed and allow flexible combinations of characters. A misspelled name may still appear as a superficially valid character sequence without triggering obvious syntactic or semantic anomalies, making context-based detection and correction difficult.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Homophonic Name Error	
Input	降雨（jiàng yǔ）喜欢画画。
Translate	Rainfall likes drawing.
Correct	江宇（jiāng yǔ）喜欢画画。
Translate	Jiang Yu likes drawing.
Near-phonetic Name Error	
Input	尘埃（chén āi）喜欢画画。
Translate	Dust likes drawing.
Correct	程艾（chéng ài）喜欢画画。
Translate	Cheng Ai likes drawing.
Orthographic Name Error	
Input	休息（xiūxī）喜欢画画。
Translate	Rest likes drawing.
Correct	林息（lín xī）喜欢画画。
Translate	Lin Xi likes drawing.

Figure 1: Examples of Chinese personal name spelling errors.

Personal-name errors in Chinese input methods are particularly common. Modern Chinese input editors often employ fuzzy pinyin matching(Zhang & Yuan, 2025) and frequency-based candidate ranking to improve typing efficiency, enabling them to retrieve high-frequency candidates with similar pronunciations even when input is incomplete or slightly inaccurate. However, for low-frequency target names, a common homophonic or near-phonetic word may be ranked above the intended name and selected by mistake. Fig 1 presents three representative phonological and orthographic errors in Chinese names, where red and blue characters indicate incorrect and correct characters, respectively. In the homophonic example, the target name '*Jiang Yu*' is replaced by the common word '*rainfall*'. In the near-phonetic case, '*Cheng Ai*' is replaced by '*dust*'. In the orthographic example, the target name '*Lin Xi*' is replaced by '*rest*'. These examples demonstrate that Chinese personal-name errors not only stem from phonetic and visual confusion at the character level (Wang et al., 2026a;b) but are also closely related to candidate ranking, name sparsity, and weak contextual constraints.

Existing personal-name correction methods typically decompose the correction process into two stages: candidate acquisition(Hou et al., 2025) and candidate selection(Jiang et al., 2024). They first identify potentially incorrect name spans or retrieve candidate names through name-span detection or name-list matching, and then evaluate the candidates using similarity features or contextual information. Udupa & Kumar (2010) and Varol & Bayrak (2012) match input names with a predefined name list to identify similar or standardized name forms. Yang et al. (2019) employ a bidirectional entity-boundary detection model to extract potentially misspelled name spans from text and rank the resulting candidates based on contextual semantic information. Singla et al. (2022) use a lightweight sequence-to-sequence model to generate name spellings from automatic speech recognition hypotheses.

Although this candidate retrieval–candidate selection paradigm explicitly leverages name boundaries and candidate-name resources, its performance largely depends on acquisition accuracy and resource coverage. If suspicious spans are not correctly identified or if the target name is not retrieved, subsequent correction cannot recover from upstream errors. The model may also fail when the target name is absent from the candidate set or when insufficient contextual evidence is available for validation. Fig 2 illustrates this error propagation. The acquired candidate name is not a real name but rather a substring formed across the word boundary between '*queen*' and '*beautiful*'.The candidate acquisition stage incorrectly treats the substring as a name candidate, and the candidate-selection stage further changes it to '*Wang Meimei*', thereby corrupting a sentence that should have remained unchanged. Therefore, the challenge in Chinese personal-name correction lies not only in the candidate selection itself, but also in the candidate acquisition's reliance on clear name boundaries and predefined name lists.

The dataset provides direct supervision for personal-name error patterns. However, our preliminary experiments show that fine-tuning a CSC model solely on name-oriented data is insufficient: al-

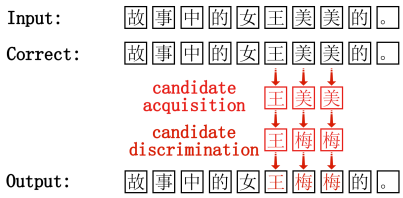


Figure 2: Illustration of the error accumulation effect in the candidate acquisition–candidate selection paradigm.

though it substantially improves target-name correction, it degrades performance on general CSC benchmarks. Name-specific training does more than just add new knowledge; it also shifts the model's correction preferences toward the distribution of target names. The core challenge is to introduce target-name correction capability while retaining the original general-domain behavior of the model as much as possible.

To address these limitations, we no longer treat name-boundary detection or candidate-name retrieval as separate prerequisite steps. Instead, we treat personal-name errors as a special type of character substitution in CSC, model them jointly with general spelling errors, and propose a Dual-Branch Chinese Spelling Correction Model (DCSC) for target Chinese personal names. Given a sentence containing spelling errors, the model directly predicts its corrected sequence at the character level, without first identifying which spans correspond to personal names or retrieving candidate entities from external name resources. DCSC enhances existing CSC models from both data and architecture perspectives. At the data level, an Online Correction Dataset of Chinese Personal Names (CPN-Corr) for pre-specified target names is constructed by the substitution of personal pronouns and erroneous-name samples generation, which includes the target names and their homophonic, near-phonetic, and orthographic variants. At the architecture level, we design a Dual-Branch Gated Fusion (DBGF) module, which separately models name-specific and general correction features.

At the architectural level, the personal-name branch and the general branch generate two residual representations, which are dynamically combined via context-dependent gating and injected into the backbone representation through scaled residual connections.

Our main contributions are as follows:

(1) We integrate Chinese personal-name correction into a unified character-level CSC framework and propose the DCSC method for target Chinese personal names, which can simultaneously handle name errors and general text errors while reducing reliance on explicit name recognition and external name resources.

(2) We propose an on-demand training-data generation scheme for Chinese personal name correction. This method constructs target-name samples through personal-pronoun replacement and homophonic, near-phonetic, and orthographic perturbations, alleviating the scarcity of name-related samples in existing CSC datasets.

(3) We introduce a lightweight dual-branch gated fusion module that coordinates name-specific features with general correction features through a personal-name branch, a general branch, and a gating fusion mechanism. This module improves the accuracy of target-name correction while maintaining the performance of general CSC.

2 Related Work

2.1 Chinese Spelling Correction

Early CSC systems relied on handcrafted rules, dictionaries, confusion sets, and statistical language models, typically decomposing error correction into error detection, candidate generation, and candidate ranking. Yu & Li (2014) used a character-level n-gram language model to identify low-probability positions and generated candidates based on phonological and visual similarity. Chang et al. (2015) combined rule-based and statistical methods and designed correction strategies tailored

to words of different lengths. While these methods are interpretable, they heavily depend on manually designed rules and the coverage of confusion sets, which limits their ability to handle diverse errors in complex contexts.

With the rise of neural networks, CSC has shifted toward end-to-end character-level sequence labeling, where models predict corrected characters at each input position(Xu et al., 2025). Han et al. (2019) applied Bi-LSTM to character-wise labeling and incorporated pinyin and stroke features to enhance detection of phonologically and visually similar errors. Pre-trained language models have further enriched contextual representations, and BERT-based models (Devlin et al., 2019) became a major foundation for recent CSC research. Soft-Masked BERT (Zhang et al., 2020) places a detection network before the correction network to improve awareness of error positions. MDCSpell (Zhu et al., 2022) models detection and correction in parallel and fuses detection information later, avoiding the loss of phonological and visual cues caused by directly masking erroneous characters. ReaLiSe (Xu et al., 2021) separately encodes semantic, phonological, and glyph-image information and integrates them through selective fusion. SCOPE (Li et al., 2022) introduces character-pronunciation prediction as an auxiliary task and leverages fine-grained pinyin modeling to improve near-phonetic correction. NamBert (Zhang et al., 2025) further explores the role of multi-modal features in CSC and applies non-aligned late fusion across semantic, pinyin, and glyph branches, providing richer information to the prediction layer. Tang et al. (2026) proposed PPMEP, a plug-and-play maximum-entropy framework that penalizes low-entropy predictions to alleviate overconfident errors in PLM-based CSC.

Large language models (LLMs) have also been introduced to CSC(Liu et al., 2025b; Zhou et al., 2024; Wu et al., 2025), despite their autoregressive generation paradigm not naturally satisfying the character-level and length-preserving constraint of the task. C-LLM (Li et al., 2024) strengthens character-wise alignment between source and target sentences through character-level tokenization. Liu et al. (2024) proposed ARM, a training-free alignment-and-replacement module that aligns LLM outputs and combines them with predictions from existing CSC models, mitigating length variation, format instability, and overcorrection. Qiao et al. (2025) injected the local correction distribution from a BERT-based small model into the LLM decoding process, combining precise local corrections with stronger language modeling capabilities without fine-tuning the LLM. Lin et al. (2026) proposed CEC-Zero, a zero-supervision reinforcement learning framework that synthesizes errorful inputs from clean text and computes cluster-consensus rewards via semantic similarity and candidate agreement, enabling LLM self-correction without human labels. Li et al. (2026) introduced CIF, a two-stage framework that decouples error recognition and revision via multimodal detection and chain-of-thought-based correction.

2.2 Personal-Name Correction

Research on English personal-name correction mainly involves name matching, normalization, and variant generation. Udupa & Kumar (2010) used hashing to map similar names to nearby binary codes, improving candidate retrieval and matching at scale. Varol & Bayrak (2012) combined edit distance, phonetic codes, and statistical information to generate normalized candidates for misspelled names. Sukharev et al. (2014) formalized name-variant generation for genealogical record retrieval as character-level machine translation, leveraging user search logs and language models to generate candidates. For spoken-name collection, Singla et al. (2022) proposed a lightweight sequence-to-sequence model capable of generating name spellings from ASR output.

Chinese personal-name correction typically follows the candidate acquisition–candidate selection paradigm. For news texts, Yang et al. (2019) first localize potentially misspelled name spans with an entity-boundary model and then rank candidate names based on spelling similarity, contextual semantic similarity, and lexical association. Several public patents adopt similar engineering pipelines. Zhao et al. (2022) first identify incorrect names and their candidates, then determine the correction result using candidate-text semantic similarity, name similarity, and positional consistency. Chen et al. (2023) annotate erroneous name entities in speech transcripts with a Chinese entity recognizer and combine a name-prediction model with a precompiled name lexicon. Yang (2023) verifies and corrects names in domain-specific texts by matching contextual relationships between names and other named entities and querying knowledge graphs. Xu et al. (2023) construct candidate sets from common-name lists and outputs of Chinese lexical-analysis tools, then detect potential spelling errors using phonological and orthographic code similarity. Although these methods ex-

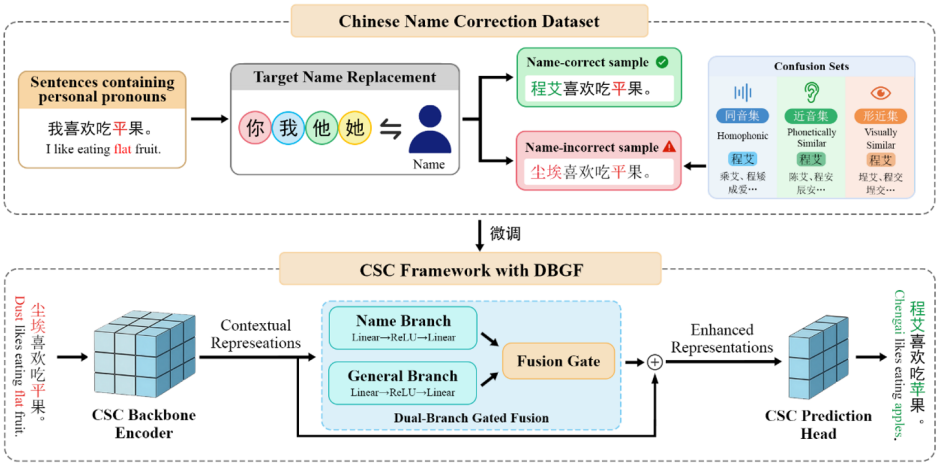


Figure 3: Overall framework of the proposed method.

plicitly leverage name boundaries, candidate sets, and domain knowledge, they remain sensitive to candidate-acquisition errors and limited coverage of external lexicons and knowledge graphs.

3 METHOD

To reduce reliance on explicit name-boundary detection, candidate-name sets, and external knowledge resources, we propose Dual-branch Chinese Spelling Correction Model (DCSC) for target Chinese personal names, integrating target-name errors into a unified CSC framework. Since specialization for target names may conflict with general CSC capabilities, our framework combines a data-construction procedure with architectural enhancement, as shown in Fig 3. First, we construct homophonic, near-phonetic, and orthographic confusion sets to simulate common name misspellings. Then, we restructure the existing CSC corpora by replacing personal pronouns with target names and perturbing those names with the confusion sets. The training split of CPN-Corr is then used to fine-tune the CSC models. Through online dataset construction, we provide supervision for target names and their easily confused variants, enabling the model to directly learn character-level mappings from erroneous sentences to corrected ones. Additionally, DBGF is inserted between the backbone encoder and the prediction layer, where the encoder generates contextual representations. The two branches and fusion gate within the DBGF module generate enhanced features. Finally, the prediction layer outputs the corrected text.

3.1 TASK DEFINITION

Given a prespecified set of target personal names, we first construct name-oriented correction data and adapt the model. During inference, the model takes an input sentence $X = \{x_1, \ldots, x_n\}$ and generates an output sentence $Y = \{y_1, \ldots, y_n\}$ of the same length. Here, x_i is the input character at position i and y_i is its target character. If $x_i = y_i$, the position remains unchanged; otherwise, the model should replace x_i with y_i. After adapting to the target names, the model directly handles both target-name errors and general text errors within the same CSC framework.

To generate realistic personal-name misspellings, we construct a candidate confusion set for each character in the target name. For a target character c, its confusion set is defined as:

$$C(c) = C_{hom}(c) \cup C_{pho}(c) \cup C_{gly}(c), \tag{1}$$

where $C_{hom}(c)$, $C_{pho}(c)$ and $C_{gly}(c)$ denote the homophonic, near-phonetic, and orthographic confusion sets, respectively.

The homophonic confusion set simulates erroneous selection among candidates with identical pronunciations in pinyin-based input. For a target character, we collect characters that share any standard Mandarin pronunciation with it, including both tone-consistent and tone-different cases. For

polyphonic characters, they are included as candidates if any of their pronunciations matches one of the target character's pronunciations, thereby improving coverage of realistic input scenarios.

The near-phonetic confusion set models substitutions that are phonetically similar but not identical. We select near-phonetic candidates from two complementary perspectives. First, Levenshtein edit distance (Levenshtein et al., 1966) measures string-level differences between pinyin forms. Second, similarities in initials and the internal structure of finals, such as the main vowel, are used to assess phonetic proximity.

The orthographic confusion set models misspellings caused by visually similar glyph structures. Candidate characters are rendered as standardized glyph images, and the structural similarity index measure (Wang et al., 2004) is used to compare the target-character image with each candidate image. For each target character, the top-K most similar characters are retained to form the orthographic confusion set.

3.2 Online Correction Dataset for Chinese Personal Names

Existing CSC datasets primarily focus on general spelling errors, and the included name examples are insufficient for models to learn target names and their common misspellings. To address this limitation, we reconstructed and expanded the SIGHAN datasets (Wu et al., 2013; Yu et al., 2014; Tseng et al., 2015) and Wang271K (Wang et al., 2018) to construct Chinese spelling correction data centered on specific target names. During data construction, we preserved as much as possible the syntactic structures, contextual semantics, and distribution of general spelling errors from the original corpora, enabling the model to learn both personal-name and general spelling corrections within a unified CSC framework.

We first select sentences containing personal pronouns from the SIGHAN datasets and Wang271K. We then apply part-of-speech tagging and rule-based filtering to remove cases where the target pronoun appears in idioms or fixed expressions. Since personal pronouns and personal names typically serve similar referential functions, replacing pronouns with a target name largely preserves the syntactic and semantic coherence of the original sentence. The resulting sentences are retained as correct-name examples. Common spelling errors already present in the source sentences are preserved so that the model remains exposed to ordinary CSC errors while learning target-name representations.

We next perturb the target names at the character level using the homophonic, near-phonetic, and orthographic confusion sets to generate erroneous-name examples. The final dataset includes correct-name examples and three types of name errors. The correct examples reduce the risk of unnecessary modifications to a valid target name, whereas the three error types simulate common sources of name misspellings, including input-method candidate errors, minor pronunciation deviations, and visual confusions. This construction preserves the general-error distribution from the original CSC corpora while explicitly introducing both correct target names and their erroneous variants, thereby providing direct supervision for personal-name correction.

3.3 Dual-Branch Gated Fusion

When the CSC model is fine-tuned only on the name-correction training data, it simultaneously learns target-name patterns and retained general correction patterns, but its backbone representation may become biased toward the distribution of target names. To alleviate this bias, we insert a Dual-Branch Gated Fusion (DBGF) module between the backbone encoder and the prediction layer. DBGF leaves the backbone architecture unchanged and adds two lightweight feed-forward residual branches on top of its output, adaptively combining name-specific features and general correction features through gating mechanisms.

Let the contextual representation produced by the CSC backbone encoder be:

$$H = [h_1; h_2; \ldots; h_n] \in \mathbb{R}^{n \times d}, h_i \in \mathbb{R}^d, \tag{2}$$

where n is the sequence length and d is the hidden dimension.

Given H, DBGF applies a personal-name branch and a general branch to produce two residual representations:

$$R^{name} = F_{name}(H), R^{gen} = F_{gen}(H), \tag{3}$$

where R^{name} and R^{gen} are the residual features generated by the personal-name and general branches, respectively, and $F_{name}(\cdot)$ and $F_{gen}(\cdot)$ denote the corresponding feed-forward transformations.

The personal-name branch models error patterns associated with target names and their homophonic, near-phonetic, and orthographic variants, whereas the general branch preserves and supplements information required for ordinary spelling correction. Both branches operate on a shared backbone representation but learn residual information with different emphases, thereby providing partially separated modeling paths for name-specific and general features, and reducing interference within a single parameter transformation.

To dynamically control the contribution of the two branches at each character position, DBGF computes a gate from the backbone representation:

$$G = \sigma(HW_g + b_g), \tag{4}$$

where σ is the sigmoid function, W_g and b_g are trainable gating parameters, and G is the fusion weight.

A larger gate value assigns more weight to the residual information from the personal-name branch, whereas a smaller value relies more on the complementary representations provided by the general branch. The two residual representations are fused as:

$$R = G \odot R^{name} + (1 - G) \odot R^{gen}, \tag{5}$$

where $\odot$ denotes element-wise multiplication and R is the fused residual representation.

Finally, DBGF injects the fused representation into the original backbone features through a scaled residual connection:

$$H' = H + \alpha R, \tag{6}$$

where H' is the enhanced hidden representation and α is a learnable scaling parameter that controls the strength of the additional residual features.

DBGF is designed to supplement rather than replace the backbone's contextual encoding capability. The personal-name branch increases sensitivity to target-name error patterns, the general branch preserves the stable features required for ordinary text correction, and the gate performs context-dependent fusion between them. Because the module relies only on the contextual representations produced by the backbone, it requires no external name dictionary, knowledge graph, or explicit name-boundary annotation. Therefore, it can be attached as a lightweight plug-in between the backbone encoder and the prediction layer of different CSC architectures.

4 Experiments

4.1 Experimental Setup

For general-domain evaluation, we use SIGHAN13 (Wu et al., 2013), SIGHAN14 (Yu et al., 2014), and SIGHAN15 (Tseng et al., 2015) to assess correction of common character and word errors. For the personal-name setting, we evaluate on CPN-Corr, which contains five target names, each with 3,000 examples, totaling 15,000 examples. The training, validation, and test splits contain 7,500, 2,500, and 5,000 examples respectively.

To evaluate whether the proposed method applies to different CSC architectures, we conduct experiments with MDCSpell, SCOPE, and NamBert. A model reproduced with its original architecture, without name-correction training data or DBGF, is referred to as the base model. Its counterpart fine-tuned on the name-correction training split and augmented with DBGF is referred to as the enhanced model. In the following tables, a model name without "(Ours)" denotes the base model, whereas "model (Ours)" denotes the enhanced variant.

Following the standard CSC evaluation, we report precision (Pre), recall (Rec), and F1 score for both detection and correction. Detection evaluates whether the model correctly identifies erroneous positions, whereas correction further requires generating the correct target character at each detected position.

4.2 Performance Comparison

Table 1 reports sentence-level results of three baseline CSC models and their enhanced variants on the SIGHAN benchmark and CPN-Corr. MDCSpell, SCOPE, and NamBert represent the base models, whereas MDCSpell(Ours), SCOPE(Ours), and NamBert(Ours) denote the variants fine-tuned on the name-correction training split and equipped with DBGF. D-P, D-R, and D-F stand for detection precision, recall, and F1, respectively; C-P, C-R, and C-F represent correction precision, recall, and F1, respectively.

Table 1: Sentence-level results of different CSC models on the SIGHAN datasets and CPN-Corr.

Dataset	Model	Detection-level			Correction-level		
		D-P	D-R	D-F	C-P	C-R	C-F
SIGHAN13	MDCSpell	86.40	80.56	83.38	84.62	78.90	81.66
	MDCSpell (Ours)	85.38	79.52	82.35	83.59	77.86	80.62
	SCOPE	86.67	81.05	83.77	85.79	80.23	82.92
	SCOPE (Ours)	85.31	**80.12**	82.63	84.21	79.09	81.57
	NamBert	87.86	81.98	84.82	86.87	81.05	83.86
	NamBert (Ours)	**87.36**	**81.87**	**84.53**	**86.15**	**80.74**	**83.36**
SIGHAN14	MDCSpell	64.97	68.85	66.85	63.70	67.50	65.55
	MDCSpell (Ours)	63.20	67.69	65.37	61.40	65.77	63.51
	SCOPE	67.45	71.73	69.52	65.64	69.81	67.66
	SCOPE (Ours)	65.44	68.46	66.92	63.42	66.35	64.85
	NamBert	66.67	73.08	69.72	65.61	71.92	68.62
	NamBert (Ours)	**65.83**	70.77	68.21	64.58	69.42	66.91
SIGHAN15	MDCSpell	78.09	81.70	79.86	75.97	79.48	77.69
	MDCSpell (Ours)	**77.13**	80.41	78.73	**75.18**	78.37	76.74
	SCOPE	80.21	83.92	82.02	78.09	81.70	79.86
	SCOPE (Ours)	**79.68**	82.62	**81.13**	**77.72**	80.59	**79.13**
	NamBert	77.11	85.95	81.29	76.45	85.21	80.59
	NamBert (Ours)	**78.88**	83.55	**81.15**	**77.49**	82.07	**79.71**
CPN-Corr	MDCSpell	42.24	40.93	41.58	39.06	37.85	38.45
	MDCSpell (Ours)	**94.20**	**94.08**	**94.14**	89.59	89.48	89.54
	SCOPE	45.00	43.97	44.48	38.16	37.28	37.72
	SCOPE (Ours)	**95.12**	**95.06**	**95.10**	**94.38**	**94.26**	**94.33**
	NamBert	37.22	35.85	36.52	36.08	34.75	35.40
	NamBert (Ours)	**94.29**	**93.62**	**93.95**	**91.80**	**91.14**	**91.47**

On the SIGHAN benchmark, the enhanced models generally maintain performance comparable to their corresponding base models. Performance remains relatively stable on SIGHAN13 and SIGHAN15. NamBert(Ours) is particularly stable: compared to NamBert, its detection F1 decreases by 0.29% and 0.14%, while correction F1 decreases by 0.50% and 0.88%, respectively. The fluctuations are more pronounced on SIGHAN14, with all three models experiencing a drop of over one percentage point in both detection and correction F1 scores. Overall, introducing name-oriented training data and DBGF does not lead to significant degradation in general correction performance.

The three enhanced models perform strongly on CPN-Corr. MDCSpell(Ours), SCOPE(Ours), and NamBert(Ours) achieve detection F1 scores of 94.14%, 95.10%, and 93.95%, respectively, indicat-

ing accurate localization of target-name errors. SCOPE(Ours) attains the highest correction F1 score at 94.33%, followed by NamBert(Ours) at 91.47% and MDCSpell(Ours) at 89.54%. All detection and correction scores are substantially higher than those of the corresponding base models. These results indicate that the constructed data and DBGF effectively support the learning of mappings between target names and their confusable variants.

4.3 Ablation Study of DBGF

To assess the effect of DBGF, Table 2 compares the enhanced models with their counterparts without DBGF on SIGHAN15 and CPN-Corr. MDCSpell(Ours), SCOPE(Ours), and NamBert(Ours) denote the corresponding base models fine-tuned on CPN-Corr with DBGF incorporated. C-P, C-R, and C-F represent correction precision, recall, and F1, respectively. The "w/o DBGF" variants are fine-tuned on the same name-correction training data but without DBGF.

Table 2: Ablation results of DBGF.

Model	SIGHAN15			CPN-Corr		
	C-P	C-R	C-F	C-P	C-R	C-F
MDCSpell (Ours)	75.18	78.37	76.74	89.59	89.48	89.54
w/o DBGF	75.71	73.75	74.72	89.91	89.34	89.62
SCOPE (Ours)	77.72	80.59	79.13	94.38	94.26	94.33
w/o DBGF	77.22	80.22	78.69	93.73	93.64	93.68
NamBert (Ours)	77.49	82.07	79.71	91.80	91.14	91.47
w/o DBGF	75.00	82.07	78.38	90.44	89.87	90.15

DBGF primarily mitigates the capability shift induced by name-specific training. Without DBGF, the correction F1 decreases on SIGHAN15 for all three backbones. Fine-tuning only on name-oriented data biases the model toward the target-name distribution and weakens its general CSC performance. Compared to the corresponding w/o DBGF variants, incorporating DBGF improves performance on SIGHAN15, indicating that the module partially restores general correction behavior. Meanwhile, the enhanced models retain strong performance on CPN-Corr: SCOPE(Ours) and NamBert(Ours) further improve name correction, while MDCSpell(Ours) shows only a 0.08% difference compared to its w/o DBGF counterpart. Thus, DBGF generally improves the balance between target-name specialization and general correction, without sacrificing one capability for another.

5 Conclusion

In this paper, we study Chinese personal-name correction under the challenges of low corpus frequency, weak semantic constraints, and frequent phonological or orthographic confusion. Instead of relying on explicit name recognition or external candidate resources, we incorporate target-name errors into a unified character-level CSC framework and propose a dual-branch Chinese spelling correction model for target Chinese personal names. We first select sentences from existing CSC corpora that contain replaceable pronouns and substitute these pronouns with the target names. Since pronouns and names typically serve similar referential functions, the replacement largely preserves the syntactic and semantic naturalness of the original sentences. Then, the target names are perturbed with phonological and orthographic confusion sets to generate realistic erroneous-name examples, thereby constructing the online CPN-Corr dataset. In addition, we design a dual-branch gated fusion module to coordinate name-specific and general correction. DCSC reduces error propagation from candidate acquisition to subsequent candidate selection while preserving the advantages of CSC models in contextual representation learning and character-level substitution prediction.

References

Tao-Hsing Chang, Hsueh-Chih Chen, and Cheng-Han Yang. Introduction to a proofreading tool for chinese spelling check task of sighan-8. In *Proceedings of the eighth SIGHAN workshop on*

Chinese language processing, pp. 50–55, 2015.

Wei Chen, Shaohui Feng, and Jianye Zhang. A method for building a personal name correction model and computer device. China, CN114818669B, 2023.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186, 2019.

Zijia Han, Chengguo Lv, Qiansheng Wang, and Guohong Fu. Chinese spelling check based on sequence labeling. In *2019 International Conference on Asian Language Processing (IALP)*, pp. 373–378. IEEE, 2019.

Zhengxu Hou, Bingren Yan, Zhibo Yang, Qiang Zhou, Luxi Xing, Chen Zhang, and Xingsheng Zhang. Revisiting the knowledge recall and selection in chinese spelling correction. In *2025 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. IEEE, 2025.

Liming Jiang, Xingfa Shen, Qingbiao Zhao, and Jian Yao. Mlsl-spell: Chinese spelling check based on multi-label annotation. *Applied Sciences*, 14(6):2541, 2024.

Vladimir I Levenshtein et al. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady*, 10(8):707–710, 1966.

Fangfang Li, Jikun Dong, Cheng Huang, Weidai Xia, and Xingliang Mao. Cif: A two-stage cognitively-inspired framework for chinese spelling correction. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 19427–19431. IEEE, 2026.

Jiahao Li, Quan Wang, Zhendong Mao, Junbo Guo, Yanyan Yang, and Yongdong Zhang. Improving chinese spelling check by character pronunciation prediction: The effects of adaptivity and granularity. In *Proceedings of the 2022 conference on empirical methods in natural language processing*, pp. 4275–4286, 2022.

Kunting Li, Yong Hu, Liang He, Fandong Meng, and Jie Zhou. C-llm: learn to check chinese spelling errors character by character. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 5944–5957, 2024.

Yinghao Li, Heyan Huang, Baojun Wang, and Yang Gao. Drmspell: dynamically reweighting multimodality for chinese spelling correction. *Frontiers of Information Technology & Electronic Engineering*, 26(3):354–366, 2025.

Zhiming Lin, Kai Zhao, Sophie Zhang, Peilai Yu, and Canran Xiao. Cec-zero: Zero-supervision character error correction with self-generated rewards. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 23612–23620, 2026.

Changchun Liu, Kai Zhang, Junzhe Jiang, Zirui Liu, Hanqing Tao, Min Gao, and Enhong Chen. Arm: An alignment-and-replacement module for chinese spelling check based on llms. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 10156–10168, 2024.

Linfeng Liu, Hongqiu Wu, and Hai Zhao. Driving chinese spelling correction from a fine-grained perspective. In *Proceedings of the 31st International Conference on Computational Linguistics*, pp. 10727–10737, 2025a.

Linfeng Liu, Hongqiu Wu, and Hai Zhao. Evolving chinese spelling correction with corrector-verifier collaboration. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 29010–29016, 2025b.

Ziheng Qiao, Houquan Zhou, and Zhenghua Li. Mixture of small and large models for chinese spelling check. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 28298–28311, 2025.

Karan Singla, Shahab Jalalvand, Yeon-Jun Kim, Ryan Price, Daniel Pressel, and Srinivas Bangalore. Seq-2-seq based refinement of asr output for spoken name capture. *arXiv preprint arXiv:2203.15833*, 2022.

Jeffrey Sukharev, Leonid Zhukov, and Alexandrin Popescul. Learning alternative name spellings. *arXiv preprint arXiv:1405.2048*, 2014.

Jinsong Tang, Jun Ren, Ruizhang Huang, and Lina Ren. Ppmep: An plug-and-play chinese spelling correction method based on the maximum entropy principle. *Journal of Chinese Information Processing*, 40(2):163–172, 2026. ISSN 1003-0077.

Yuen-Hsien Tseng, Lung-Hao Lee, Li-Ping Chang, and Hsin-Hsi Chen. Introduction to sighan 2015 bake-off for chinese spelling check. In *Proceedings of the Eighth SIGHAN Workshop on Chinese Language Processing*, pp. 32–37, 2015.

Raghavendra Udupa and Shaishav Kumar. Hashing-based approaches to spelling correction of personal names. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pp. 1256–1265, 2010.

Cihan Varol and Coskun Bayrak. Hybrid matching algorithm for personal names. *Journal of Data and Information Quality (JDIQ)*, 3(4):1–18, 2012.

Dingmin Wang, Yan Song, Jing Li, Jialong Han, and Haisong Zhang. A hybrid approach to automatic corpus generation for chinese spelling check. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, pp. 2517–2527, 2018.

Shuo Wang, Lei Jiang, Nannan Sun, and Yanbing Liu. Rlcsc: Reinforcement learning enhanced chinese spelling correction with glyph-phonetic similarity. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3106–3110. IEEE, 2026a.

Shuo Wang, Chaodong Tong, and Lei Jiang. Hcl-csc: Hierarchical contrastive learning with ids-aware character similarity for chinese spelling correction. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 17212–17216. IEEE, 2026b.

Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

Haiming Wu, Hanqing Zhang, Richeng Xuan, and Dawei Song. Bi-dcspell: a bi-directional detector-corrector interactive framework for chinese spelling check. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pp. 3974–3984, 2024.

Haiming Wu, Zhinie Nie, Songkun Ji, and Dawei Song. Dra: A dual retrieval architecture for domain chinese spelling check. In *CCF International Conference on Natural Language Processing and Chinese Computing*, pp. 250–262. Springer, 2025.

Shih-Hung Wu, Chao-Lin Liu, and Lung-Hao Lee. Chinese spelling check evaluation at sighan bake-off 2013. In *Proceedings of the seventh SIGHAN workshop on Chinese language processing*, pp. 35–42, 2013.

Heng-Da Xu, Zhongli Li, Qingyu Zhou, Chao Li, Zizhen Wang, Yunbo Cao, He-Yan Huang, and Xian-Ling Mao. Read, listen, and see: Leveraging multimodal information helps chinese spell checking. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pp. 716–728, 2021.

Jianbing Xu, Weijin Fei, Jun Li, and Lijuan Zhang. A method, system, device and storage medium for personal name recognition and correction. China, CN116341543B, 2023.

Mingying Xu, Jie Liu, Kui Peng, and Zhen Li. Chinese spelling correction based on long short-term memory network-enhanced transformer and dynamic adaptive weighted multi-task learning. *Natural Language Processing*, 31(5):1265–1284, 2025.

Yue Yang, Ruizhang Huang, Qin Wei, Yanping Chen, and Yongbin Qin. A method for correction of person names in news based on contextual semantics. *Journal of University of Electronic Science and Technology of China*, 48(6):809–814, 2019. DOI: 10.3969/j.issn.1001-0548.2019.06.002.

Zizhao Yang. A context-semantics-based method, system, medium and terminal for personal name correction in a specific domain. China, CN115630632A, 2023.

Junjie Yu and Zhenghua Li. Chinese spelling error detection and correction based on language model, pronunciation, and shape. In *Proceedings of The Third CIPS-SIGHAN Joint Conference on Chinese Language Processing*, pp. 220–223, 2014.

Liang-Chih Yu, Lung-Hao Lee, Yuen-Hsien Tseng, and Hsin-Hsi Chen. Overview of sighan 2014 bake-off for chinese spelling check. In *Proceedings of The Third CIPS-SIGHAN Joint Conference on Chinese Language Processing*, pp. 126–132, 2014.

Lei Zhang, Zhetan Liu, Qianxi Yan, and Xiaodong Liu. Psif: Phonetic–semantic and long–short information fusion for chinese spelling correction. *Applied Sciences*, 16(5):2440, 2026.

Shaohua Zhang, Haoran Huang, Jicong Liu, and Hang Li. Spelling error correction with soft-masked bert. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, pp. 882–890, 2020.

Xiaowu Zhang, Hongfei Zhao, Jingyi Hou, and Zhijie Liu. Unveiling the impact of multimodal features on chinese spelling correction: From analysis to design. *arXiv preprint arXiv:2504.07661*, 2025.

Zijian Zhang and Jinfeng Yuan. Decopy: detect and correct with pinyin for chinese spelling correction. *Royal Society Open Science*, 12(9):250426, 2025.

Honghong Zhao, Baoxin Wang, Dayong Wu, and Zhigang Chen. Personal name correction method, apparatus, electronic device and storage medium. China, CN113947073A, 2022.

Houquan Zhou, Zhenghua Li, Bo Zhang, Chen Li, Shaopeng Lai, Ji Zhang, Fei Huang, and Min Zhang. A simple yet effective training-free prompt-free approach to chinese spelling correction based on large language models. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 17446–17467, 2024.

Chenxi Zhu, Ziqiang Ying, Boyu Zhang, and Feng Mao. Mdcspell: A multi-task detector-corrector framework for chinese spelling correction. In *Findings of the association for computational linguistics: ACL 2022*, pp. 1244–1253, 2022.