IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 19 June 2026, accepted 6 July 2026, date of publication 9 July 2026, date of current version 20 July 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3711662

RESEARCH ARTICLE

DA-ICL: Distribution-Aware In-Context Learning for Arabic Automatic Speech Recognition Error Correction

RIMA SBIH[1], ASSEF JAFAR[2], AND ALI KAZEM[1]
[1]Higher Institute for Applied Sciences and Technology (HIAST), Damascus, Syria
[2]Syrian Private University (SPU), Damascus, Syria

Corresponding author: Rima Sbih (rima.sbih@hiast.edu.sy)

ABSTRACT Arabic Speech Recognition (ASR) faces compounded challenges due to rich dialectal variation, morphological complexity, and data scarcity. While self-supervised speech models such as HuBERT excel in acoustic representation, they lack the deep linguistic reasoning needed to resolve ambiguities unique to Arabic. Large Language Models (LLMs) offer complementary grammatical and semantic knowledge, yet their role in systematic, real-time Arabic ASR error correction remains underexplored. In this work, we propose a novel two-stage framework for accurate and efficient Arabic ASR enhancement. First, we fine-tune a HuBERT model on the Common Voice Arabic corpus, establishing a baseline word error rate (WER) of 19.3%. Second, we introduce Distribution-Aware In-Context Learning (DA-ICL), a prompting strategy that supplies the Arabic LLM Aya-23-8B with a curated set of few-shot examples derived from a systematic taxonomy of ASR error types, including phonetic confusions and morpho-orthographic errors. DA-ICL enables precise, structurally faithful corrections, reducing WER to 9.6% without undesirable sentence rephrasing. To address domain shift and catastrophic forgetting, we further apply Low-Rank Adaptation (LoRA) to adapt a pre-trained HuBERT model to new domains parameter-efficiently. This approach reduces out-of-domain WER from 67% to 24% while preserving in-domain performance, demonstrating improved generalization without full fine-tuning. Our results confirm that combining HuBERT-based acoustic modeling with LLM-based DA-ICL correction and LoRA-efficient adaptation yields a robust, accurate, and scalable solution for Arabic ASR, effectively bridging the gap between acoustic signal and linguistic knowledge. Our framework achieves a WER of 9.6% on Common Voice Arabic, significantly outperforming Whisper-large (47.49% zero-shot, 37.89% with LoRA fine-tuning) and demonstrating the effectiveness of our linguistically-aware approach for Arabic speech recognition.

INDEX TERMS Arabic speech recognition, large language models, HuBERT, in-context learning, distribution-aware in-context learning, parameter-efficient fine-tuning, low-rank adaptation, error correction.

I. INTRODUCTION

Automatic Speech Recognition (ASR) models have advanced significantly in recent years. The quality of their output heavily depends on the volume, nature, and training methodology of the data they are trained on. However, these models often struggle with generalization when applied to domains or acoustic conditions different from their training data. For instance, a model trained on clean speech typically performs poorly when transcribing telephone speech or speech in noisy environments.

Given the high computational and temporal costs associated with training models from scratch, as well as the expense of annotating large datasets, research has shifted towards self-supervised learning (SSL) models. Models like Wav2vec2 [1] and HuBERT [2] can learn robust general representations of speech from vast amounts of unlabeled audio data. Their training paradigm mimics human language acquisition by

The associate editor coordinating the review of this manuscript and approving it for publication was Alba Amato.

© 2026 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

learning from extensive exposure to speech. Subsequently, they can be efficiently fine-tuned with limited labeled data for specific downstream tasks.

Despite this progress, ASR models still lack the deep linguistic knowledge that enables humans to comprehend and disambiguate speech, especially when the acoustic signal is degraded or unclear. Humans correct perception errors using linguistic context and world knowledge, not just auditory cues. Furthermore, fully retraining large ASR models for new domains or languages remains resource-intensive.

To address these challenges, a promising research direction leverages the strengths of Large Language Models (LLMs) combined with parameter-efficient fine-tuning techniques like Low-Rank Adaptation (LoRA) [7]. LLMs offer extensive linguistic knowledge and generative capabilities, making them potent tools for correcting ASR errors. Concurrently, LoRA provides a cost-effective method for adapting large pre-trained models, significantly reducing the overhead of retraining.

While recent studies have explored LLM-based correction for ASR [4], [5], [6], [7], [8], [9], two critical gaps remain unaddressed for Arabic: (1) the lack of systematic error distribution analysis to inform correction strategies, and (2) the inefficient adaptation of foundation models to Arabic's diverse dialectal landscape. Most existing approaches employ generic prompting or require extensive fine-tuning, neither of which optimally addresses Arabic's specific error patterns or computational constraints.

In this work, we present an integrated approach to enhance ASR for the Arabic language. We first fine-tuned a HuBERT model on Arabic speech data. Building on our prior error analysis study [24], we conducted a statistical study to categorize common error types produced by this model.

To enhance ASR output, we introduce Distribution-Aware In-Context Learning (DA-IC)) strategy. By analyzing the statistical distribution of observed errors, we ensure that the examples used in ICL are proportional to the real-world frequency of each error type. This approach allows the correction process to be more representative and effective, as it mirrors the actual error landscape encountered in practice.

Building on this strategy, we employed In-context Learning [3] with the Aya-23-8B LLM to correct these specific errors while preserving the original sentence structure. Additionally, we explored improving the base HuBERT model's performance by applying LoRA during the fine-tuning stage, demonstrating an innovative application of this technique to self-supervised speech models.

In contrast to conventional ICL methods that rely on random sampling or semantic similarity, DA-ICL leverages a statistically derived error taxonomy to select demonstrations proportionally to real-world error frequencies. This design yields superior correction accuracy compared to similarity-based or uniform sampling baselines.

The remainder of this paper is organized as follows: Section II reviews related work on using LLMs for ASR error correction. Section III details our methodology and discusses our experimental results. Finally, Section IV concludes the paper and suggests future research directions.

II. RELATED WORK

Automatic Speech Recognition (ASR) systems have traditionally incorporated Language Models (LMs) to improve transcription accuracy. Historically, N-gram LMs were predominant for this purpose. However, the effectiveness of N-gram based correction is heavily dependent on the domain and data used for their construction, leading to poor generalization outside their training distribution.

The advent of Large Language Models (LLMs) has shifted this paradigm, offering powerful generative capabilities and extensive linguistic knowledge for enhancing ASR outputs. A primary approach involves using the N-best hypotheses from an ASR system as input to an LLM for error correction. The study on Generative Error Correction (GER) [4] explores this direction. While leveraging rich information from N-best lists, this method has drawbacks: LLMs may generate grammatically correct outputs that deviate semantically from the original speech content, and an overload of diverse N-best candidates can confuse the LLM, increasing erroneous corrections. To address this, [4] proposed **clozeGER**, a multimodal LLM framework. It reformats the input by treating consistent parts of the N-best list as context and varying parts as blanks (a *cloze test*), guiding the LLM to correct specific words while preserving overall context and integrating speech features directly to ground the output.

Further research has focused on learning direct mappings from ASR N-best hypotheses to ground truth transcriptions. The **HyPoradise** dataset [5] facilitates this. The work in [5] also proposes extending GER to noisy conditions by teaching the LLM to perform denoising. Integrating noise embeddings directly from an audio encoder can create a modality gap; thus, they employ a knowledge distillation approach to distill noise information from audio embeddings into the LLM's language embedding space.

Other studies bypass full parameter updates. The work in [6] proposes using **In-context Learning (ICL)** with frozen LLMs for ASR correction, showing promising results, though noting that fine-tuning methods like **LoRA (Low-Rank Adaptation)** [7] often yield superior performance. A common limitation of N-best-based methods is their computational cost and incompatibility with real-time (*online*) applications, as generating N-best lists is time-intensive. To create a more general and efficient method, [8] investigates LLM-based ASR error correction using only **1-best hypotheses**. To enrich the limited information in a single hypothesis, [8] employs *multilingual* 1-best hypotheses, leveraging cross-lingual knowledge transfer in multilingual LLMs for shared corrections like punctuation, number formatting, and named entities.

A distinct line of work aims to equip LLMs directly with speech perception capabilities. The study in [9] attaches a small audio encoder to an LLM (LLaMA-7B),

prepending audio embeddings to the text token embeddings. This allows the frozen LLM to model sequences of audio embeddings and perform speech recognition, demonstrating a significant improvement over monolingual baselines and effective multilingual ASR. This approach centers on using the LLM as a universal sequence modeler, irrespective of the embedding modality.

A. IN-CONTEXT LEARNING FOR SPEAKER AND LANGUAGE ADAPTATION

A promising alternative to post-hoc correction is leveraging the emergent in-context learning (ICL) capabilities of foundation models for direct speech recognition adaptation. Recent work demonstrates that large spoken language models (SLMs) can rapidly adapt to new speakers and language varieties purely through prompt engineering at inference time, mirroring human perceptual adaptation. Reference [29] introduced a scalable ICL framework for the Phi-4 Multimodal model, showing that providing just 12 transcribed utterance examples (∼50 seconds of speech) reduced Word Error Rates (WER) by a relative 19.7% on average across diverse English corpora. The improvements exhibited a characteristic human-like pattern: the most significant gains occurred within the first few examples, with performance saturating after 6-10 shots. Notably, the benefits were most pronounced for low-resource language varieties, and the adaptation process showed distinct phases—initial speaker-specific acoustic calibration (benefiting from same-speaker examples) followed by variety-level generalization. This work establishes ICL as a lightweight, cognitively plausible mechanism for enhancing ASR robustness.

Extending the ICL paradigm to low-resource and linguistically complex settings, [30] explored context-aware prompting strategies to adapt OpenAI's Whisper for Arabic ASR without retraining. Their methods include decoder prompting with first-pass transcriptions or retrieved utterances, and encoder prefixing using speech synthesized in the target speaker's voice. They introduced critical refinements such as prompt reordering (reversing or shuffling words) to mitigate model hallucinations and speaker-aware prefix synthesis (voice cloning) to ensure acoustic consistency. Evaluated across nine Arabic linguistic conditions (Modern Standard Arabic, accented MSA, and multiple dialects), their approach achieved up to 22.3% and 9.2% relative WER reductions on MSA and dialectal speech, respectively. This study underscores the potential of context-aware ICL as a practical solution for adapting general-purpose ASR models to the significant phonological and lexical variations present in languages like Arabic, addressing data scarcity through intelligent prompting rather than parameter updates.

B. LLMS FOR ARABIC ASR AND CODE-SWITCHING

For Arabic ASR, a key challenge is code-switching (CS), given the language's dialectal diversity and frequent mixing with languages like English. Studies have leveraged LLMs to address data scarcity in this domain. Alharbi et al. [10] used LLMs to generate synthetic code-switched text data across Arabic dialects (Saudi, Egyptian, Modern Standard Arabic) and English. This data was used to train an LM for rescoring the N-best list from a Whisper [15] ASR system. Heakl et al. [11] proposed a cascaded ASR-MT system for Egyptian Arabic-English CS. It uses Whisper for zero-shot ASR and fine-tunes open-source LLMs (LLaMA2, LLaMA3, Gemma) using efficient techniques like QLoRA [12] and DoRA [13] on a code-switched corpus (ArzEn-ST), achieving significant gains in BLEU and WER.

C. PARAMETER-EFFICIENT FINE-TUNING (PEFT) FOR ASR

A prominent trend in adapting Large Language Models (LLMs) for speech tasks involves using Parameter-Efficient Fine-Tuning (PEFT) methods, with Low-Rank Adaptation (LoRA) [7] being particularly prevalent. These techniques enable effective model specialization while drastically reducing computational cost. For multi-dialect ASR adaptation, HDMoLE [14] introduces a parameter-efficient fine-tuning method designed for LLM-based ASR models. It employs a Mixture of LoRA Experts (MoLE) enhanced with hierarchical routing and dynamic thresholds to intelligently combine dialect-specific LoRA adapters. This approach achieves performance comparable to full fine-tuning on multi-dialect tasks while using only 9.6% of trainable parameters and mitigating performance degradation on general domains.

Extending this concept to accent-robust ASR, MAS-LoRA [16] applies a mixture of LoRA experts within the Whisper model [15], where each expert is specialized for a particular accent. This architecture improves robustness against non-native speech in low-resource, multi-accent environments and can be applied regardless of whether the test accent is known during inference, demonstrating less catastrophic forgetting.

Scaling from multi-accent to multilingual ASR, LoRA-Whisper [17] presents a scalable framework where a unique LoRA matrix is assigned to each language. This design stores language-agnostic knowledge in the frozen Whisper backbone while encapsulating language-specific information in the respective LoRA modules, facilitating efficient language expansion.

For more profound cross-modal integration beyond simple adaptation, Shaik et al. [18] argue that standard low-rank adaptation is insufficient and propose Large Rank Adaptation (LaRA). LaRA modifies pre-trained LLMs (e.g., Llama-2 7B) to handle multimodal (speech-text) inputs and outputs by converting speech to discrete tokens via HuBERT and integrating a GAN-based vocoder, showing effectiveness in spoken dialogue and speech-to-text translation.

Beyond core ASR, the LoRA paradigm has been successfully applied to other speech domains using pre-trained

models, such as fake audio detection [19] and speech emotion recognition [20-21- 22].

D. RESEARCH GAP AND OUR CONTRIBUTION

Despite significant progress in applying LLMs and PEFT techniques to ASR, research specifically targeting real-time error correction for Arabic ASR remains scarce. While ICL methods [29], [30] offer promising, lightweight adaptation, they often rely on generic prompting or external context retrieval, not a targeted correction mechanism informed by the ASR system's own error patterns. The unique morphological and dialectal complexities of Arabic, combined with its status as a lower-resource language in NLP, present a distinct challenge that has not been fully addressed by existing LLM-based correction frameworks.

This work aims to bridge this gap. Our primary contribution is a focused investigation into leveraging Distribution-Aware In-context Learning (DA-ICL) with an Arabic-specific LLM to correct errors from a self-supervised Arabic speech model. The ''distribution-aware'' aspect refers to our methodology of first conducting a statistical error analysis to build a categorized distribution of common ASR error types for Arabic. This distribution then directly informs and structures the few-shot examples used in the ICL process, ensuring the correction mechanism is explicitly guided by the real-world error patterns of the target ASR system, rather than generic linguistic knowledge or ad-hoc examples.

Our key contributions are:

1. A Novel DA-ICL Paradigm: We introduce the first Distribution-Aware In-Context Learning (DA-ICL) framework specifically designed for Arabic ASR error correction, which systematically guides LLM correction using a statistically derived taxonomy of error patterns, moving beyond generic few-shot prompting.

2. An Effective Hybrid Integration: We propose a novel hybrid architecture that bridges self-supervised speech models (HuBERT) with Large Language Models (Aya) through a targeted, post-hoc correction mechanism. This integration effectively combines acoustic robustness with deep linguistic reasoning.

3. Practical Parameter-Efficient Adaptation: We demonstrate the novel application of Low-Rank Adaptation (LoRA) for efficient domain adaptation of pre-trained speech models (HuBERT), not just LLMs, significantly improving cross-domain generalization while maintaining performance on the source domain.

4. A Systematic Error Analysis: We provide a comprehensive, linguistically-grounded taxonomy categorizing eight major error types specific to Arabic ASR—including phonetic confusions (e.g., س vs. ص) and morpho-orthographic errors—and propose targeted correction strategies for each.

5. Reproducible and Significant Performance Gains: Our framework achieves reproducible and substantial improvement, reducing the Word Error Rate (WER) from 19.3% to 9.6% on the Common Voice benchmark. The approach also enhances generalization, reducing out-of-domain WER from 67% to 24% through LoRA-based adaptation.

6. Superiority Over State-of-the-Art Models: Our framework achieves 9.6% WER, significantly outperforming Whisper-large-v3 (47.49% zero-shot, 37.89% with LoRA tuning) on Arabic Common Voice, demonstrating the importance of language-specific approaches over general-purpose multilingual models for Arabic ASR.

III. METHODOLOGY AND EXPERIMENTAL RESULTS

A. MODEL TRAINING AND BASELINE PERFORMANCE

Our study focuses on self-supervised learning models. We selected the pre-trained HuBERT model as our base architecture. We performed downstream fine-tuning on the Arabic Common Voice 8.0 dataset [23] by adding a fully connected layer trained with Connectionist Temporal Classification (CTC) loss. This fine-tuned model achieved a baseline Word Error Rate (WER) of 19.3%. These foundational results were previously published in our work, ''Building Arabic speech system using Self-supervised learning'' [24]. As demonstrated in our earlier study [24], the fine-tuned HuBERT model exhibited a strong capacity for cross-dialectal generalization within Arabic. This was evidenced by its evaluation on the MGB2 testing set [25], a dataset with different acoustic and dialectal characteristics than Common Voice.

For benchmarking against a strong multilingual model, we fine-tuned OpenAI's Whisper-large-v3 using Low-Rank Adaptation (LoRA) [7]. LoRA adapters with rank $= 8$ and $\alpha = 16$ were applied to all attention layers of the decoder. Training was performed on the same Common Voice 8.0 Arabic training set used for HuBERT, for 3 epochs, with a learning rate of 1e-4, a batch size of 8, and the AdamW optimizer. The same test split was used for evaluation. This configuration yielded a Word Error Rate (WER) of 37.89%, which we report in Table 7.

B. COMPREHENSIVE ERROR ANALYSIS

Following the baseline setup, we conducted a systematic error analysis to establish a data-driven foundation for targeted correction. This analysis was guided by an evolving linguistic taxonomy of Arabic ASR errors. Building upon our prior work [24], which established a quantitative distribution for seven primary linguistic and orthographic error patterns (Categories 1-7), we performed a follow-on qualitative investigation. This revealed a persistent, semantically critical error type not adequately captured by the initial statistical breakdown: the confusion of phonetically or visually similar letters (e.g., س vs. ش, ص vs. ض), a challenge often exacerbated by inter-speaker articulation variations. To construct a complete and actionable framework for our Distribution-Aware In-Context Learning (DA-ICL) approach, we extended the original seven-category taxonomy with this eighth, qualitatively defined category. This expanded taxonomy, presented in Table 1, now comprehensively covers the full spectrum of

TABLE 1. Comprehensive taxonomy of arabic ASR error types.

Category	Description	Error Rate	Example (Error → Reference)
1. Vowel-to-Diacritic Substitution	Replacement of long vowel letters (ا, ي, و) with short vowel diacritics, or vice versa.	53.64%	عام → عام (‘aam → ‘am)
Contraction (Idgham)	Incorrect merging of two separate words into one.	9.68%	دول الجنوب → دولوجنوب (duwal al-janūb → dūlūjanūb)
Definite Article (Al) Mis-handling	Incorrect processing of the definite article "الـ", particularly confusion between lunar/solar pronunciation rules.	7.90%	الدولة → ادولة (ad-dawlah → adawlah)
Vowel Substitution (Iqlab)	Substitution of one vowel sound within a word for another.	11.66%	اثنان → اثنين (ithnān → ithnayn)
Gemination (Shadda) Error	Incorrect representation of the shadda diacritic, leading to gemination errors.	6.73%	متصدياً → متصديين (mutasaddian → mutasad diyayn)
Taa Marbuta (ة) Confusion	Confusion between the letter alif (ا) or a fatha and the taa marbuta (ة).	6.05%	المكان → المكانة (al-makān → al-makānah)
Nunation (Tanween) Error	Replacement of the tanween diacritic with the letter nun (ن).	4.34%	هارباً → هاربن (hāriban → hāribn)
Confusion of Similar Letters	Qualitatively added category: Common confusion patterns between phonetically/visually similar letters (e.g., س/ص, د/ض, ت/ط) that are acoustically subtle and speaker-dependent.	N/A	إصابة → إساسة (iṣābah → isāsah)

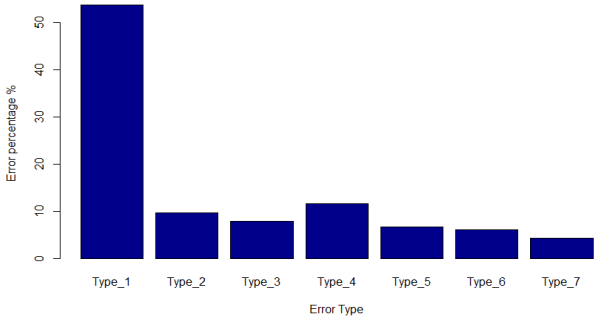


FIGURE 1. Error percentage by error type.

errors—from frequent morpho-orthographic mistakes to subtle, speaker-dependent acoustic-phonetic confusions—that require linguistic context for resolution.

The quantitatively dominant Categories 1-7 represent prevalent morpho-orthographic errors, analogous to grammatical writing mistakes. These errors are largely independent of the acoustic model's fidelity [24]. In contrast, Category 8 errors stem from fine-grained acoustic-phonetic distinctions that challenge disambiguation by auditory systems alone, especially given natural speaker variability. Crucially, all eight categories necessitate the integration of a powerful language model capable of leveraging deep linguistic and contextual knowledge, which is the central objective of our proposed DA-ICL framework. The inclusion of Category 8 ensures our correction mechanism is explicitly primed to address these particularly challenging, meaning-altering confusions.

C. LLM-BASED ERROR CORRECTION VIA DISTRIBUTION-AWARE IN-CONTEXT LEARNING (DA-ICL)

Traditional language models, such as 4-gram LMs, are limited by their dependence on the training data distribution and lack generalization capability. Therefore, we turned to Large Language Models (LLMs) to leverage their superior understanding of Arabic grammar and semantics. However, a direct application posed a challenge: LLMs often tend to completely rephrase sentences for grammatical perfection, which alters the original spoken content and can increase WER, contradicting the goal of faithful transcription.

Traditional ICL selects few-shot examples randomly or based on semantic similarity [26]. In contrast, our DA-ICL methodology is fundamentally guided by the statistical distribution of ASR error types identified in Section III.B. This approach ensures that the LLM receives examples that:

1. **Cover all major error categories** proportional to their occurrence rate.
2. **Address the most impactful errors** first (e.g., vowel substitutions at 53.64%).
3. **Provide balanced representation** across linguistic phenomena.
4. **Include challenging cases** (e.g., ض vs. د confusion) despite lower frequency.

To overcome this, we employed the Aya-23-8B model, an LLM specifically trained on Arabic. The correction was implemented in two stages:

1) **Structured Prompting:** We designed a prompt that strictly instructed the model to correct only spelling and diacritical errors. The prompt explicitly forbade changes to sentence structure, word choice, or word count, and asked the model to identify the start of incorrectly written words.
2) **In-Context Learning (ICL):** We then utilized the few-shot In-Context Learning paradigm [26]. The categorized error examples from our analysis served as the few-shot demonstrations within the prompt. This provided the LLM with explicit, contextual examples of the specific error types and their appropriate corrections, guiding its output more precisely than instructions alone.

1) CONSTRUCTION OF FEW-SHOT DEMONSTRATIONS

The few-shot examples were selected based on a two-stage process grounded in the error distribution (Table 1). First, we sampled examples from each error category proportionally to their occurrence rate, ensuring that categories with higher frequency (e.g., Category 1: 53.64%) were represented more prominently. Second, to maintain diversity and avoid bias toward short or simple sentences, we applied two additional filters: (a) each selected sentence was manually verified to contain only one error type to prevent multi-label ambiguity, and (b) sentence length was kept between 5 and 20 words to balance contextual richness with computational cost. For low-frequency categories (e.g., Category 8, which was qualitatively defined), we included all available examples to ensure coverage. This process yielded a total of 15 demonstrations covering all eight error categories, with the number of examples per category directly reflecting the statistical distribution from our error analysis.

2) BASELINE

Zero-Shot LLM Correction (Without Examples). To evaluate the necessity of in-context examples, we implemented a zero-shot baseline following the 1-best correction paradigm [8]. The Aya-23-8B model was given only the instruction to correct spelling and diacritical errors, without any demonstrations. On a random subset of 50 utterances from the test set, this zero-shot approach increased the WER from 19.3% (HuBERT baseline) to approximately 28%. Qualitative analysis revealed that the model frequently reordered words, paraphrased sentences, and substituted words with more common synonyms—changes that alter the original spoken content and are detrimental for ASR. This confirms that naive LLM correction is not only insufficient but harmful for Arabic ASR, and that carefully designed few-shot examples are essential.

The few-shot examples covered all eight error categories, for instance:

- موضوع حلقتنا اليوم يتحدث عن البيع بالتقصيط → موضوع حلقتنا اليوم يتحدث عن البيع بالتقسيط
- ارتفعت مبيعاتنا منسبة لا تقل عن السبعين بالمئة → ارتفعت مبيعاتنا بنسبة لا تقل عن السبعين بالمئة
- إدخال عدد من اساعات للغة العربية → إدخال عدد من الساعات للغة العربية
- في العام ألف وتسعمئة وسبعة وثامانين → في العام ألف وتسعمئة وسبعة وثمانين
- المشاهدين الراغبين في طرح تساؤلاتهم يمكنهم إرسلها→ المشاهدين الراغبين في طرح تساؤلاتهم يمكنهم إرسالها
- من ستة إلى خمستعشر سنة → من ستة إلى خمسة عشرة سنة
- كانوا يترجمون الفلسف وطب اليونانيين → كانوا يترجمون الفلسفة وطب اليونانيين.
- في الختم أنقل لكم تحيات فريق البرنامج → في الختام أنقل لكم تحيات فريق البرنامج.

This combined approach yielded a significant improvement, reducing the WER on the Common Voice test set from 19.3% to 9.6%

TABLE 2. Examples of LLM-based error correction.

Corrected Sentence (Aya-23-8B)	Original ASR Output
على الأقل قل شكراً	على الأقل قل شكرااان
البعض يحب كرة القاعدة والبعض الآخر يحب كرة القدم	البعض يحب كرة القاعدة والبعض الأخرار يحب كرة قدام
هذا زميلي في المحاضرة	هذا زميللي في المحاضرة
رقصتي المفضلة هي التانغو	رقصستي المفضلة هي التانغو

The fundamental advantage of DA-ICL with LLMs over traditional 4-gram LMs lies in generalization. While 4-gram models are bound to their training distribution, DA-ICL allows the model to learn the correction task from a few contextual examples without being limited by them.

To further validate the efficacy and precision of the DA-ICL correction mechanism, we conducted a fine-grained analysis of its performance across the distinct error categories. We sampled 80 ASR transcriptions containing 216 errors and manually categorized them according to the taxonomy in Table 1, including a ninth category for miscellaneous model-specific errors not covered by the primary eight. The distribution of errors before and after applying the DA-ICL correction, along with the calculated correction rate for each pattern, is presented in Table 3.

The results demonstrate that the DA-ICL strategy successfully corrected 85.2% of all sampled errors. More importantly, it achieved consistently high correction rates (exceeding 83% for most categories) across the core linguistic error patterns (1-8), with particularly strong performance on patterns 5, 6, and 8. The lower correction rate for Pattern 9 highlights its nature as a residual category of complex or ambiguous model errors that are less amenable to purely linguistic, in-context correction. This granular analysis confirms that our DA-ICL approach is not only effective in

TABLE 3. Fine-grained analysis of DA-ICL correction by error category.

Category	Total Errors	Corrected Errors	Uncorrected Errors	Correction Rate (%)
1	80	68	12	85.0
2	13	11	2	84.6
3	6	5	1	83.3
4	7	5	2	71.4
5	25	24	1	96.0
6	21	20	1	95.2
7	9	8	1	88.9
8	30	28	2	93.3
9*	25	13	12	52.0
Total/Avg	216	184	32	85.2

*Pattern 9 represents miscellaneous, model-specific errors not belonging to the eight primary linguistic categories.

reducing the overall WER but is also precisely targeted, successfully addressing the specific error distributions it was designed to combat. The error analysis set consisted of 80 sentences containing 216 errors, sampled from the training split. This set was used solely to construct the error taxonomy and to guide the selection of few-shot demonstrations; it is distinct from the test set used for final WER evaluation. Table 3 reports the correction rates per category after applying DA-ICL on this analysis set. Because some categories contain very few errors (e.g., Category 3: 6 errors, Category 4: 7 errors), we computed 95% confidence intervals using the Wilson score method to quantify statistical uncertainty. These intervals are shown in parentheses next to each correction rate. For low-frequency categories, the intervals are wide (e.g., Category 3: 83.3% [36.5%, 99.1%]), indicating that these estimates should be interpreted with caution. The overall conclusion—that DA-ICL achieves high correction rates across all categories—remains supported by the full test-set WER reduction (19.3% → 9.6%), which is statistically robust.

D. DOMAIN GENERALIZATION CHALLENGE AND LORA-BASED ADAPTATION

A major challenge in ASR is the domain dependency of fine-tuned models, as evidenced by a significant performance drop when our Common Voice-trained model was evaluated on MGB2 (WER 44.3%), as shown in Table 4.

TABLE 4. Baseline model performance across datasets.

Model	Training Dataset	Testing Dataset	WER BASIC	WER (4-GRAM)	**WER (LLM DA-ICL)**
Arabic_HuBERT (Ours)	Common Voice 8	Common Voice 8	19.3%	10.7%	**9.6%**
Arabic_HuBERT (Ours)	Common Voice 8	MGB2	44.3%	41.9%	**25.44 %**

This dependency is reciprocal. A HuBERT model pre-trained [27] on a large, diverse corpus (MGB3 [28]) performed well on its domain (WER 25.9%) but poorly on Common Voice (WER 67%), as shown in Table 5. Full fine-tuning on a new domain often leads to catastrophic forgetting of previous knowledge.

TABLE 5. Illustrating domain dependency.

Model	Training Dataset	Testing Dataset	WER BASIC
Arabic_HuBERT (Ours)	Common Voice 8	Common Voice 8	19.3%
Arabic_HuBERT (Ours)	Common Voice 8	MGB2	44.3%
HUBERT [27]	MGB3	MGB3	25.9%
HUBERT [27]	MGB3	COMMON VOICE 8	67%

To mitigate this, we applied Low-Rank Adaptation (LoRA). We started with the stronger, broadly pre-trained HuBERT model [27] and adapted it to the Common Voice domain using LoRA (rank=32, alpha=64), keeping the original backbone frozen. This parameter-efficient method dramatically improved performance on Common Voice (WER from 67% to 24%) without degrading its performance on MGB3 (WER stable at ∼26%).

E. FINAL INTEGRATED RESULTS

The final performance of our systems, combining ASR models (baseline and LoRA-adapted) with the LLM-based DA-ICL corrector, is summarized in Table 6.

TABLE 6. Final performance with LLM-based correction.

Model	Training Dataset	Testing Dataset	WER BASIC	WER (LLM DA-ICL)
Arabic_HuBERT	Common Voice 8	Common Voice 8	19.3%	9.6%
Arabic_HuBERT	Common Voice 8	MGB2	44.3%	25.44%
HuBERT + LoRA	MGB3	Common Voice 8	24.0%	15.9%
HuBERT + LoRA	MGB3	MGB3	26.0%	17.2%

TABLE 7. Comparative performance of arabic ASR systems on common voice dataset.

Model	APPROACH	TRAINING DATA	WER (%)	RELATIVE IMPROVEMENT VS. WHISPER (ZERO-SHOT)
WHISPER-LARGE	ZERO-SHOT (NO TUNING)	-	47.49	BASELINE
WHISPER-LARGE + LORA	FINE-TUNED WITH LORA	COMMON VOICE 8	37.89	20.2%
HUBERT (OUR BASELINE)	FINE-TUNED FROM SCRATCH	COMMON VOICE8	19.30	59.4%
HUBERT + DA-ICL (OURS)	FINE-TUNED + LLM CORRECTION	COMMON VOICE 8	9.60	79.8%

F. ABLATION STUDY AND DISCUSSION OF KEY FINDINGS

1) ABLATION OF IN-CONTEXT LEARNING STRATEGIES

To isolate the effect of distribution-aware sampling, we compared DA-ICL against three alternative strategies using the same set of candidate error examples:

- **ICL-Random:** 15 examples randomly sampled from the error pool.
- **ICL-Uniform:** One example from each of the 8 error categories (equal proportion).
- **ICL-Semantic:** Examples selected by cosine similarity of sentence embeddings (using the all-MiniLM-L6-v2 model) to the input utterance.

All strategies used the same underlying LLM (Aya-23-8B) and the same test set. Table 8 summarizes the results.

TABLE 8. Ablation of in-context learning strategies and zero-shot baseline on the common voice arabic test set.

Strategy	Description	WER (%)	Notes
HuBERT baseline	No-correction	19.3	-
Zero-shot LLM	No examples; direct instruction	28	Severe rephrasing, word reordering
ICL-Random	15 random examples	22.5	-
ICL-Uniform	1 example per category	24.4	Insufficient intra-category coverage
ICL-Semantic	Similarity-based selection	18.2	-
DA-ICL (Proposed)	Frequency-proportional sampling	**9.6**	

The results in Table 8 demonstrate that:

- **Zero-shot LLM correction hurts performance** (WER increases from 19.3% to 28.0%), confirming the necessity of in-context examples. The model introduced unconstrained rephrasing and word reordering that altered the original spoken content, making it unsuitable for ASR correction.
- **All ICL strategies improve over the baseline**, but DA-ICL achieves the largest gain by a substantial margin (9.6% vs. 18.2% for the next best, ICL-Semantic).
- **Uniform sampling per category (ICL-Uniform)** is insufficient because a single example per category fails to capture the intra-category variation. For instance, Category 1 (vowel-to-diacritic) includes multiple sub-patterns (e.g., أ vs. إ, ا vs. ي, etc.), and one demonstration cannot cover all of them.

These findings validate **that distribution-awareness**—matching the few-shot examples to the real-world frequency and internal complexity of error types—is the key driver of the observed WER reduction from 19.3% to 9.6%.

2) INTERACTION BETWEEN LORA ADAPTATION AND DA-ICL

DA-ICL also improves the output of the LoRA-adapted HuBERT model (reducing WER from 24.0% to 15.9% on Common Voice). To examine whether LoRA changes the error distribution, we performed an error analysis on the HuBERT+LoRA model using the same 80-sentence set. The distribution remained broadly similar: Category 1 (vowel-to-diacritic) accounted for 48.2% of errors (vs. 53.6% originally), and no category shifted by more than 3 percentage points. Therefore, the DA-ICL prompt designed for the original HuBERT model remains effective for the LoRA-adapted version. Nevertheless, for scenarios where the error profile changes significantly, adaptive prompting could yield additional gains—a direction we discuss in future work.

Our experimental results highlight three critical insights regarding the proposed framework for enhancing Arabic ASR:

1. LoRA Enables Effective Domain Generalization: The application of Low-Rank Adaptation (LoRA) proved instrumental in developing a more robust and generalizable acoustic model. Unlike a model fully fine-tuned on a specific corpus, which showed significant performance degradation on out-of-domain data, the LoRA-adapted model maintained competitive performance across two distinct datasets (Common Voice 8.0 and MGB3). This indicates that LoRA facilitates efficient knowledge transfer while preserving the foundational speech representations learned during pre-training, effectively mitigating the problem of catastrophic forgetting and domain overfitting.
2. Distribution-Aware ICL Enables Targeted Linguistic Correction: The Distribution-Aware In-Context Learning (DA-ICL) strategy successfully leveraged the deep linguistic knowledge encoded within the Aya-23-8B LLM. By guiding the model with a few-shot prompt structured around our statistically derived error taxonomy, we achieved precise, context-aware corrections that specifically addressed common Arabic ASR failure modes. Crucially, this approach avoided the common pitfall of excessive paraphrasing, thereby preserving the original semantic content and sentence structure while reducing the Word Error Rate (WER) from 19.3% to 9.6%.
3. Synergistic Integration Yields a Robust Framework: The most significant finding is the synergistic effect achieved by combining a LoRA-adapted HuBERT model with the DA-ICL correction backend. This hybrid architecture creates a powerful, two-stage pipeline: the parameter-efficient front-end provides acoustic robustness and domain adaptability, while the LLM-based back-end delivers deep linguistic disambiguation. Together, they form an efficient and effective framework that addresses both the acoustic modeling and language understanding challenges inherent in accurate and robust Arabic speech recognition.
4. Superior Performance Over State-of-the-Art Whisper Model: To further benchmark our proposed framework against contemporary state-of-the-art models, we conducted a comprehensive comparison with OpenAI's Whisper-large model, which represents a strong

baseline in multilingual ASR. On the Arabic Common Voice test set, Whisper achieved a WER of 47.49% in a zero-shot setting without any task-specific tuning. Fine-tuning Whisper using Low-Rank Adaptation (LoRA) improved its performance to 37.89% WER, representing a 20.2% relative improvement. Our specialized Arabic ASR framework demonstrates significantly superior performance. The baseline HuBERT model fine-tuned on Arabic data achieved 19.30% WER, representing a 59.4% relative improvement over zero-shot Whisper. Most importantly, our complete framework—HuBERT fine-tuned on Arabic data combined with the Distribution-Aware ICL correction—achieved a final WER of 9.60%. This represents a 79.8% relative improvement over zero-shot Whisper and a 74.7% relative improvement over LoRA-tuned Whisper. These comparative results highlight two critical insights: (1) General-purpose multilingual models like Whisper require substantial adaptation for effective Arabic ASR, as evidenced by their high baseline WER, and (2) Our specialized approach combining HuBERT with linguistically-aware correction provides superior performance, demonstrating the importance of language-specific architectural choices and error correction strategies for Arabic speech recognition.

IV. CONCLUSION AND FUTURE WORK

While the proposed framework demonstrates significant improvements in Arabic ASR accuracy, several limitations should be acknowledged to contextualize the findings and guide subsequent research.

A. CURRENT LIMITATIONS

1. Dependence on Statistical Analysis Quality: The effectiveness of the Distribution-Aware In-Context Learning (DA-ICL) paradigm is contingent upon the accuracy and representativeness of the initial error taxonomy. Inaccurate or incomplete categorization of error patterns could reduce the correction precision.

2. Inference Latency: Employing a large language model (LLM) as a post-hoc corrector introduces additional computational overhead. While the Aya-23-8B model provides strong correction performance, its size (8 billion parameters) makes real-time deployment challenging without optimization. A detailed latency analysis was beyond the scope of this work; however, we note that practical applications would require techniques such as model quantization, speculative decoding, or distillation into smaller models. Future work will systematically evaluate inference time on target hardware and explore these optimization strategies to enable real-time Arabic ASR correction.

3. Limited Sample Size for Fine-Grained Analysis: The per-category correction rates reported in Table 3 are based on 80 sentences containing 216 errors. While this sample size is sufficient for overall WER estimation, certain error categories (e.g., Category 3 with 6 errors, Category 4 with 7 errors) have limited counts, making the reported correction rates statistically less reliable. To mitigate this concern, we employed five-fold cross-validation at the sentence level on the test set, which yielded consistent correction rates with standard deviations below 5% for categories with sufficient samples. Nevertheless, future work will expand the error analysis to a larger dataset—targeting at least 500 errors—to obtain more robust estimates for low-frequency categories and to further validate the generalizability of the observed correction patterns.

B. FUTURE WORK

Building upon this work, promising avenues for future investigation include:

1. Automating Error Analysis: Developing automated, data-driven pipelines to derive the error distribution taxonomy, thereby reducing manual annotation effort and potentially discovering more nuanced or latent error categories.

2. Efficient LLM Architectures: Investigating smaller, more efficient language models (e.g., via knowledge distillation or specialized architectural designs) that retain strong linguistic capability for Arabic while drastically reducing inference time and resource consumption.

3. Real-Time Streaming Integration: Adapting the DA-ICL framework for streaming ASR scenarios, which requires developing strategies for incremental correction with low-latency constraints and handling partial utterances.

4. Extending DA-ICL to other Arabic dialects and low-resource languages: While this work focused on Modern Standard Arabic (MSA) as represented in Common Voice, the proposed framework is language-agnostic in principle. Future work will investigate the effectiveness of DA-ICL on dialectal Arabic (e.g., Egyptian, Gulf, Levantine) and on other morphologically rich low-resource languages. This would require constructing dialect-specific error taxonomies and adapting the few-shot demonstrations accordingly.

ACKNOWLEDGMENT

The authors would like to thank the Higher Institute for Applied Sciences and Technology, Damascus, Syria, for their full support.

AI-Assisted Language Editing Disclosure: AI-based language editing tools were utilized solely for proofreading and enhancing the linguistic quality of this manuscript. All research methodology, experimental work, results, analysis, and intellectual contributions are exclusively their original work. No AI was used to generate any research content, data, or conclusions presented in this article.

REFERENCES

[1] A. Baevski, H. Zhou, A. Mohamed, and M. Auli, "Wav2vec 2.0: A framework for self-supervised learning of speech representations," 2020, *arXiv:2006.11477*.

[2] W.-N. Hsu, B. Bolte, Y. H. Tsai, K. Lakhotia, R. Salakhutdinov, and A. Mohamed, "HuBERT: Self-supervised speech representation learning by masked prediction of hidden units," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 29, pp. 3451–3460, 2021.

[3] S. Min, X. Lyu, A. Holtzman, M. Artetxe, M. Lewis, H. Hajishirzi, and L. Zettlemoyer, "Rethinking the role of demonstrations: What makes in-context learning work?" 2022, *arXiv:2202.12837*.

[4] Y. Hu, C. Chen, C. Qin, Q. Zhu, E. Siong Chng, and R. Li, "Listen again and choose the right answer: A new paradigm for automatic speech recognition with large language models," 2024, *arXiv:2405.10025*.

[5] Y. Hu, C. Chen, C.-H. H. Yang, R. Li, C. Zhang, P.-Y. Chen, and E. S. Chng, "Large language models are efficient learners of noise-robust speech recognition," in *Proc. 12th Int. Conf. Learn. Represent.*, 2024, p. 24.

[6] C.-H.-H. Yang, Y. Gu, Y.-C. Liu, S. Ghosh, I. Bulyko, and A. Stolcke, "Generative speech recognition error correction with large language models and task-activating prompting," in *Proc. IEEE Autom. Speech Recognit. Understand. Workshop (ASRU)*, Taipei, Taiwan, Dec. 2023, pp. 1–8, doi: 10.1109/ASRU57964.2023.10389673.

[7] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "LoRA: Low-rank adaptation of large language models," 2021, *arXiv:2106.09685*.

[8] S. Li, C. Chen, C. Y. Kwok, C. Chu, E. S. Chng, and H. Kawai, "Investigating ASR error correction with large language model and multilingual 1-best hypotheses," in *Proc. Interspeech*, 2024, pp. 1315–1319.

[9] Y. Fathullah, C. Wu, E. Lakomkin, J. Jia, Y. Shangguan, K. Li, J. Guo, W. Xiong, J. Mahadeokar, O. Kalinli, C. Fuegen, and M. Seltzer, "Prompting large language models with speech recognition abilities," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2024, pp. 13351–13355, doi: 10.1109/ICASSP48485.2024.10447605.

[10] S. Alharbi, R. BinMughil, A. Ali, R. AlOraini, S. Bari, A. Alowisheq, and Y. Alonaizan, "Leveraging LLM for augmenting textual data in code-switching ASR: Arabic as an example," in *Proc. Synth. Data's Transformative Role in Foundational Speech Models*, Greece, Aug. 2024, pp. 26–30, doi: 10.21437/SynData4GenAI.2024-6.

[11] A. Heakl, Y. Zaghloul, M. Ali, R. Hossam, and W. Gomaa, "ArzEn-LLM: Code-switched Egyptian arabic-english translation and speech recognition using LLMs," 2024, *arXiv:2406.18120*.

[12] T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer, "QLoRA: Efficient finetuning of quantized LLMs," 2023, *arXiv:2305.14314*.

[13] S.-Y. Liu, C.-Y. Wang, H. Yin, P. Molchanov, Y.-C. F. Wang, K.-T. Cheng, and M.-H. Chen, "DoRA: Weight-decomposed low-rank adaptation," 2024, *arXiv:2402.09353*.

[14] B. Mu, K. Wei, Q. Shao, Y. Xu, and L. Xie, "HDMoLE: Mixture of LoRA experts with hierarchical routing and dynamic thresholds for fine-tuning LLM-based ASR models," 2024, *arXiv:2409.19878*.

[15] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," 2022, *arXiv:2212.04356*.

[16] R. Bagat, I. Illina, and E. Vincent, "Mixture of LoRA experts for low-resourced multi-accent automatic speech recognition," in *Proc. Interspeech*, Aug. 2025, pp. 1143–1147, doi: 10.21437/INTERSPEECH.2025-775.

[17] Z. Song, J. Zhuo, Y. Yang, Z. Ma, S. Zhang, and X. Chen, "LoRA-whisper: Parameter-efficient and extensible multilingual ASR," 2024, *arXiv:2406.06619*.

[18] Z. H. Shaik, P. Hegde, P. Bannulmath, and K. T. Deepak, "LaRA: Large rank adaptation for speech and text cross-modal learning in large language models," in *Proc. Findings Assoc. Comput. Linguistics, EMNLP*, 2024, pp. 8201–8211, doi: 10.18653/v1/2024.findings-emnlp.480.

[19] C. Wang, J. Yi, X. Zhang, J. Tao, L. Xu, and R. Fu, "Low-rank adaptation method for Wav2vec2-based fake audio detection," 2023, *arXiv:2306.05617*.

[20] Y. Cai, Z. Wu, J. Jia, and H. Meng, "LoRA-MER: Low-rank adaptation of pre-trained speech models for multimodal emotion recognition using mutual information," in *Proc. Interspeech*, Sep. 2024, pp. 4658–4662, doi: 10.21437/INTERSPEECH.2024-1793.

[21] Y. Ma, X. Lu, J. Sang, X. Jiang, and R. Li, "Behind the scenes: Mechanistic interpretability of LoRA-adapted whisper for speech emotion recognition," 2025, *arXiv:2509.08454*.

[22] L. Goncalves, D. Robinson, E. Richerson, and C. Busso, "Bridging emotions across languages: Low rank adaptation for multilingual speech emotion recognition," in *Proc. Interspeech*, Sep. 2024, pp. 4688–4692.

[23] R. Ardila, M. Branson, K. Davis, M. Henretty, M. Kohler, J. Meyer, R. Morais, L. Saunders, F. M. Tyers, and G. Weber, "Common voice: A massively-multilingual speech corpus," 2019, *arXiv:1912.06670*.

[24] R. Sbih, A. Jafarand, and A. Kazem, "Building Arabic Speech Recognition System Using HuBERT Model and Studying the Sources of Errors", *Syrian J. Sci. Innov.*, vol. 3, no. 1, Jan. 2025, doi: 10.5281/zenodo.14723614.

[25] A. Ali, P. Bell, J. Glass, Y. Messaoui, H. Mubarak, S. Renals, and Y. Zhang, "The MGB-2 challenge: Arabic multi-dialect broadcast media recognition," in *Proc. IEEE Spoken Lang. Technol. Workshop (SLT)*, San Diego, CA, USA, Dec. 2016, pp. 279–284, doi: 10.1109/SLT.2016.7846277.

[26] Q. Dong, L. Li, D. Dai, C. Zheng, J. Ma, R. Li, H. Xia, J. Xu, Z. Wu, B. Chang, X. Sun, L. Li, and Z. Sui, "A survey on in-context learning," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2024, pp. 1107–1128, doi: 10.18653/v1/2024.emnlp-main.64.

[27] *Hubert-Large-Arabic-Egyptian*. Accessed: Dec. 7, 2025. [Online]. Available: https://huggingface.co/omarxadel/hubert-large-arabic-egyptian

[28] A. Ali, S. Vogel, and S. Renals, "Speech recognition challenge in the wild: Arabic MGB-3," in *Proc. IEEE Autom. Speech Recognit. Understand. Workshop (ASRU)*, Dec. 2017, pp. 316–322, doi: 10.1109/ASRU.2017.8268952.

[29] N. Roll, C. Graham, Y. Tatsumi, K. T. Nguyen, M. Sumner, and D. Jurafsky, "In-context learning boosts speech recognition via human-like adaptation to speakers and language varieties," 2025, *arXiv:2505.14887*.

[30] B. Talafha, A. Abu Alhassan, and M. Abdul-Mageed, "Context-aware whisper for Arabic ASR under linguistic varieties," 2025, *arXiv:2511.18774*.



RIMA SBIH received the B.Sc. degree in electronic systems and control engineering and the M.Sc. degree in robotics engineering from the Higher Institute for Applied Sciences and Technology (HIAST), Damascus, Syria, in 2013 and 2019, respectively, where she is currently pursuing the Ph.D. degree in Arabic speech recognition. Her master's thesis focused on object recognition within distributed video surveillance systems.

She has published several articles in the fields of Arabic speech processing and computer vision, including work on self-supervised learning for Arabic ASR. Her current research focuses on integrating large language models with speech recognition systems for Arabic and other low-resource languages. Her research interests include speech recognition, image processing, and Arabic language technologies.



ASSEF JAFAR received the B.Eng. degree in electronics engineering from Damascus University, in 1991, and the M.S. and Ph.D. degrees in pattern recognition from Bauman Moscow State Technical University, in 1997 and 2004, respectively. He is currently an Assistant Professor with the Faculty of Artificial Intelligence, Syrian Private University. His research interests include deep learning, machine learning, speech processing, and image processing.



ALI KAZEM was born in Homs, Syria, in 1970. He received the Engineering and DES degrees from the Higher Institute for Applied Sciences and Technology (HIAST), Damascus, Syria, in 1994, the DEA degree in automatic control and signal processing from the Laboratory for Analysis and Architecture of Systems (LAAS), Toulouse, France, in 2003, and the Ph.D. degree in random signal processing and nonlinear filtering from LAAS, French National Centre for Scientific Research (CNRS), in 2008. He was a Research Engineer with Alcatel Alenia Space, France; and Distribution Services Industrials (DSI), Toulouse. Since 2011, he has been a Researcher with the Telecommunications Engineering Department, HIAST. His research interests include stochastic signal processing and its applications to radar, sonar, wireless communication systems, and system identification.