International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 05 May-2026 | Impact Factor: 3.5

Oru Koottu: A Survey on Floating UI-Based Digital Assistants for Elderly and Low-Literacy Users

Dilshana Sherin B, Shahana Sherin P K,
Aleesha P, Riya Fathima K P, Shanida T K
Department of Computer Science and Engineering
College of Engineering Thalassery
Kerala, India

Abstract—Digital services have become integral to everyday life; however, elderly and low-literacy users continue to face sig-nificant barriers in accessing and understanding modern mobile interfaces. Prior research has independently explored Optical Character Recognition (OCR), machine translation for low-resource languages, speech-based interaction, and accessibility-oriented user interface design. Despite notable advancements, existing solutions largely remain application-specific, document-centric, or require explicit user actions, limiting their effective-ness in dynamic mobile environments [1], [2].

This paper presents a comprehensive literature survey and comparative analysis of enabling technologies relevant to accessibility-driven digital assistants. Based on identified research gaps, the paper conceptually proposes *Oru Koottu*, a floating user interface-based digital assistant designed to function as a persistent, system-level accessibility layer.The proposed frame-work integrates on-screen content capture, OCR-based text extraction, contextual translation, text simplification, Malayalam voice output, speech-driven input assistance, intent-aware content classification, malicious link detection, and guided accessibility interaction to support elderly and low-literacy users without disrupting their interaction flow.

By synthesizing insights across multiple research domains, this survey highlights the need for unified, real-time assistive systems that operate across applications. The findings establish a struc-tured foundation for future research on integrated accessibility solutions aimed at bridging the digital divide and promoting inclusive access to digital services.

Index Terms—Floating user interface, optical character recog-nition, text simplification, machine translation, speech-to-text, digital accessibility

I. Introduction

The rapid expansion of digital services has transformed how individuals access information, communicate, and in-teract with essential services such as banking, healthcare, and governance. Smartphones, in particular, have become the primary gateway to these services. However, despite widespread device availability, a substantial segment of the population—especially elderly individuals and users with low or limited literacy—continues to face significant barriers in effectively using modern mobile applications [1], [3].

Most contemporary mobile interfaces are designed with assumptions of high reading proficiency, familiarity with English, and prior digital experience. Text-heavy layouts, complex navigation structures, frequent interface updates, and unfamiliar terminology often result in confusion, cognitive overload, and exclusion for elderly and low-literacy users [2]. These challenges persist even when accessibility guidelines are applied, as many standards fail to adequately address age-related cognitive decline, language barriers, and the need for contextual assistance [3].

Prior research has explored assistive technologies such as OCR, machine translation, text simplification, and speech-based interaction. However, most solutions remain application-specific and require explicit user actions such as copying text or switching applications [4], [5].

Recent advances in artificial intelligence, particularly transformer-based language models and multilingual transla-tion systems, have significantly improved translation quality and contextual understanding [6]. However, the deployment of these models in accessibility-focused systems remains lim-ited. Most AI-driven tools function as standalone applications rather than as continuous assistive layers that can dynamically interpret and support user interaction across diverse mobile contexts.

A review of accessibility-oriented interface research further reveals a lack of integrated, system-level solutions capable of providing real-time assistance across applications [1], [7]. Existing approaches often emphasize redesigning individual applications rather than augmenting user interaction at the operating-system level, thereby limiting scalability and adop-tion.

Motivated by these observations, this paper presents a literature survey and comparative analysis of enabling tech-nologies for accessibility-driven digital assistants. Based on the identified research gaps, the paper conceptually proposes *Oru Koottu*, a floating user interface-based digital assistant designed as a persistent accessibility layer that integrates on-screen content capture, contextual translation, text simplifi-cation, and speech-based assistance within the user's natural interaction flow.

The remainder of this paper is organized as follows. Section II discusses background and enabling technologies. Section III presents a structured literature review. Section IV provides a comparative analysis and identifies research gaps. Section V introduces the conceptual framework of the proposed system. Section VI outlines potential applications and use cases. Sec-tion VII discusses future research directions, and Section VIII concludes the paper.

II. Background and Enabling Technologies

A. Optical Character Recognition (OCR)

Optical Character Recognition (OCR) enables the extraction of machine-readable text from images, scanned documents, and screen content. Modern OCR techniques employ deep learning and image preprocessing approaches to improve robustness across varying fonts, layouts, and image qualities [5], [8].

For accessibility-oriented applications, OCR bridges the gap between visual content and assistive technologies. Many digital interfaces contain non-selectable text embedded within images or PDFs, making them inaccessible to traditional screen readers. OCR enables automatic text extraction for translation, simplification, and speech-based interaction, particularly benefiting elderly and low-literacy users [1].

B. NLP-Based Text Simplification

Text simplification transforms complex or technical language into simpler and more understandable forms while preserving meaning. Recent research increasingly relies on transformer-based and data-driven models trained on simplification corpora [9].

For elderly and low-literacy users, linguistic complexity often limits comprehension even after translation. NLP-based simplification improves readability by reducing sentence complexity, simplifying vocabulary, and restructuring information into more accessible forms [2].

C. Machine Translation for Regional Languages

Machine Translation (MT) converts text from a source language to a target language while preserving meaning. Neural machine translation has significantly improved multilingual translation quality, although regional and low-resource languages still face challenges related to limited training data and linguistic diversity [4], [10].

In multilingual societies, translation into regional languages is essential for inclusive digital access. Accessibility-oriented translation systems must additionally consider contextual clarity and cultural understanding to support elderly and low-literacy users effectively [6], [11].

D. Floating User Interfaces and Accessibility Overlays

Floating user interfaces provide an alternative to traditional application-centric designs. Floating interfaces remain accessible across applications and provide persistent assistance without interrupting user workflow.

For elderly and digitally inexperienced users, such interfaces reduce cognitive load by offering simplified access mechanisms and contextual guidance. Accessibility research highlights the effectiveness of persistent overlays in improving usability and reducing interaction complexity [1]–[3].

E. Speech-Based Interaction Technologies

Speech-based technologies such as Speech-to-Text (STT) and Text-to-Speech (TTS) enable voice-driven interaction with digital systems. These technologies are widely used in virtual assistants and accessibility tools [12], [13].

For elderly and low-literacy users, speech interaction reduces dependence on typing and visual navigation. Combined with OCR and translation systems, speech technologies support accessible information delivery and voice-guided interaction [7].

F. Intent Classification and Content Understanding

Intent classification identifies the semantic purpose or category of digital content using NLP and machine learning techniques. In accessibility-oriented systems, intent-aware analysis helps distinguish government notices, banking information, advertisements, spam, and suspicious content.

Recent AI-based classification models improve contextual understanding and digital safety for elderly and low-literacy users by reducing confusion regarding unfamiliar online information.

G. Malicious Link Detection and Source Reliability

Phishing attacks, fraudulent websites, and misinformation pose significant risks for digitally inexperienced users. Malicious link detection systems analyze URLs using blacklist databases, Safe Browsing APIs, domain characteristics, and machine learning techniques to identify unsafe content.

Source reliability analysis helps users avoid harmful interactions and improves trust in digital services, particularly within accessibility-focused assistive environments.

H. Accessibility APIs and Guided Interaction

Modern accessibility APIs allow assistive systems to interpret interface elements such as buttons, text fields, and navigation components. These APIs support context-aware guidance, visual highlighting, and voice-assisted navigation.

Guided interaction systems reduce cognitive overload by providing step-by-step assistance during tasks such as login, form filling, and navigation. Such approaches are particularly beneficial for elderly and first-time smartphone users.

III. Literature Review

A. OCR-Based Text Extraction and Scene Text Recognition

Prior research has extensively investigated Optical Character Recognition (OCR) techniques for extracting text from images, scanned documents, and natural scenes. Traditional OCR pipelines use preprocessing techniques such as grayscale conversion, thresholding, and skew correction, followed by recognition engines like Tesseract [5], [8]. Document-centric OCR studies report high accuracy for structured inputs such as invoices and printed documents [8].

Recent works extend OCR to natural scene text recognition for Indian and non-Latin scripts. These studies highlight challenges including complex backgrounds, illumination variations, and font diversity [5]. Deep learning models such as CNNs and CRNNs improve recognition accuracy for low-resource languages. However, most OCR systems still function as standalone components without integration into accessibility-oriented assistive systems [5], [8].

B. Machine Translation for Low-Resource and Regional Languages

Machine Translation (MT) has evolved from statistical approaches to neural machine translation (NMT) architectures. Early English-to-Malayalam systems used statistical and hybrid methods but were constrained by limited parallel corpora [10].

Recent research focuses on transformer-based NMT models with attention mechanisms for low-resource languages. Models such as MarianMT improve English–Malayalam translation quality and contextual understanding [4], [9]. Domain-specific adaptations further highlight the linguistic complexity of Dravidian languages [9], [11].

Despite these improvements, MT systems continue to face challenges related to contextual ambiguity and morphological richness. Most systems also require explicit user interaction and provide limited support for low-literacy users [4], [11].

C. Text Simplification and Summarization for Improved Comprehension

Several studies identify linguistic complexity as a major barrier to comprehension. Transformer-based models such as T5 and BART are widely used for abstractive summarization in technical and legal domains [9].

Research combining simplification with translation shows improved readability for users unfamiliar with technical terminology [9]. However, these systems are primarily evaluated in offline environments and are rarely integrated into real-time assistive applications.

D. Speech-Based Interaction Technologies

Speech-based interaction technologies, including Speech-to-Text (STT) and Text-to-Speech (TTS), support natural human–computer interaction. Traditional approaches relied on Hidden Markov Models, while modern systems increasingly use deep learning architectures such as LSTMs and transformers [12], [13].

Research on Malayalam speech recognition highlights challenges including speaker variability and noisy environments [12]. TTS systems have improved in intelligibility but still face limitations in multilingual synthesis and prosody modeling [13].

Several studies emphasize the value of speech interfaces for elderly and low-literacy users. However, most systems remain application-specific rather than functioning as integrated assistive frameworks [7].

E. Accessibility-Oriented Interfaces for Low-Literacy and Elderly Users

Research on accessibility-oriented interfaces consistently reports that text-heavy layouts and complex navigation create barriers for elderly and low-literacy users [1], [3]. Visual metaphors, icon-based interaction, and audio guidance are identified as effective alternatives [1], [7].

Systematic reviews consolidate usability principles related to pictography, audio-visual aids, and simplified interaction flows [2]. Despite these insights, most accessibility-focused solutions remain application-specific and do not support persistent cross-application assistance [2], [3].

F. Integrated AI-Based Assistive Systems and Emerging Gaps

Recent studies highlight the growing role of large multilingual AI models in improving translation quality for low-resource languages [6], [14]. Models such as NLLB-200 extend translation support to hundreds of underrepresented languages with improved contextual accuracy [6].

However, most AI-driven accessibility solutions function as standalone applications or document-level tools requiring explicit user interaction [14], [15]. Current systems rarely integrate OCR, translation, simplification, speech output, and accessibility overlays into a unified assistive framework.

G. AI-Driven Digital Safety and Accessibility Assistance

Recent accessibility systems increasingly incorporate AI-driven safety and contextual assistance mechanisms. Research on phishing detection, malicious URL identification, and interface interpretation highlights the growing importance of digital safety for vulnerable users.

Studies integrating accessibility APIs with speech guidance demonstrate that step-by-step assistance can reduce user confusion during digital workflows. Similarly, intent-aware classification models help users distinguish official notices, advertisements, and suspicious content.

Despite these advances, existing systems remain fragmented and rarely combine accessibility assistance, translation, fraud prevention, contextual guidance, and speech interaction within a unified floating assistive framework.

IV. Comparative Analysis and Research Gaps

The reviewed literature spans OCR-based text extraction, machine translation for low-resource languages, speech-based interaction technologies, and accessibility-oriented interface design. Although each domain has progressed independently, existing solutions remain significantly fragmented [1], [2].

A. Comparative Analysis of Existing Approaches

OCR-focused studies primarily improve text recognition accuracy through preprocessing, optimized OCR configurations, and deep learning-based recognition models [5], [8]. These approaches achieve strong performance for structured documents and natural scene text but generally operate as standalone pipelines without integration into user-centric assistive systems [5].

Machine Translation research for English–Malayalam and other low-resource languages has evolved from statistical approaches to transformer-based neural architectures [4], [6], [10], [11]. Models such as MarianMT and NLLB-200 improve translation quality, yet most systems still require explicit user actions such as copying text or uploading documents [2], [3].

Text simplification and summarization studies focus on reducing linguistic complexity for improved comprehension, particularly in technical and legal domains [9]. Although combining summarization with translation improves readability, these systems are mostly evaluated in offline settings rather than real-time assistive environments [9].

Speech-based interaction technologies, including Speech-to-Text (STT) and Text-to-Speech (TTS), are widely explored as accessibility enablers [12], [13]. Despite high accuracy under controlled conditions, challenges remain in handling continuous speech and noisy environments. Existing speech systems are also largely application-specific rather than persistent assistive layers [7].

Accessibility-oriented interface research consistently highlights the limitations of text-heavy and complex mobile interfaces for elderly and low-literacy users [1]–[3], [7]. Proposed solutions emphasize simplified navigation, voice guidance, and personalization, but most implementations remain confined to individual applications [2].

B. Identified Research Gaps

Based on the comparative analysis, several critical research gaps can be identified:

- **Lack of Integrated Assistive Frameworks:** Existing studies address OCR, translation, speech interaction, and accessibility as isolated components [1], [2]. Unified assistive systems integrating these technologies remain largely absent.
- **Limited Cross-Application Support:** Most solutions are application-bound and require context switching during user interaction [3]–[5].
- **Insufficient Focus on Real-Time Interaction:** Many approaches are evaluated in offline or document-based settings rather than dynamic on-screen environments [8], [9].
- **User Burden and Cognitive Overload:** Existing tools often require manual actions such as text selection or file uploads, which are unsuitable for elderly and low-literacy users [2], [3].
- **Limited Accessibility-Oriented Integration:** Few systems combine translation, simplification, speech output, and guided interaction into a persistent assistive interface [1], [7].
- **Limited Digital Safety Assistance:** Existing accessibility tools rarely integrate phishing detection, malicious URL analysis, or source reliability verification.
- **Absence of Guided Interaction Mechanisms:** Most systems do not provide contextual navigation support or step-by-step interaction guidance for elderly users.
- **Lack of Intent-Aware Accessibility:** Current assistive technologies often fail to identify the semantic intent or risk level of digital content.

These gaps motivate the need for a system-level accessibility assistant capable of capturing on-screen content, simplifying and translating information, and providing multimodal assistance through a persistent floating interface. This observation forms the conceptual foundation for the proposed *Oru Koottu* digital assistant.

V. Conceptual Framework of Oru Koottu Digital Assistant

Based on the research gaps identified in the preceding section, this paper conceptually proposes *Oru Koottu*, a system-oriented digital assistant designed to provide continuous, accessibility-focused support for elderly and low-literacy mobile users. Rather than introducing a new standalone application, the proposed framework envisions an intelligent assistive layer that operates persistently across mobile applications through a floating user interface.

A. Design Philosophy

The conceptual design of *Oru Koottu* is grounded in three core principles derived from the literature:

- **Non-intrusive assistance:** The assistant should not disrupt the user's primary task or require frequent context switching.
- **Minimal cognitive load:** Interaction should rely on simple gestures, audio guidance, and visual cues rather than text-heavy interfaces.
- **Always-available accessibility:** Assistance must be available across applications without requiring explicit user initiation for each task.

These principles directly address limitations identified in existing OCR, translation, speech, and accessibility solutions, which are often application-bound and interaction-heavy.

B. High-Level System Overview

Conceptually, *Oru Koottu* functions as an intelligent intermediary between the user and the mobile operating system. The system is envisioned as a layered framework composed of four logical components:

- **Floating Interaction Layer:** A persistent, movable on-screen element that provides entry points for assistance while allowing uninterrupted interaction with background applications.
- **Content Perception Layer:** Responsible for acquiring on-screen textual content through system-level access or visual capture mechanisms, enabling both structured and unstructured text understanding.
- **Language Intelligence Layer:** Performs text simplification, multilingual translation, and contextual interpretation using AI-based language models.
- **Accessibility and Guidance Layer:** Delivers simplified information through visual summaries, Malayalam voice narration, speech-driven interaction, UI highlighting, contextual navigation guidance, and safety alerts tailored to elderly and low-literacy users.

Fig. 1 presents the layered conceptual architecture of the proposed Oru Koottu digital assistant, highlighting the interaction among the floating assistive interface, system perception, language intelligence, and accessibility output layers. This modular abstraction allows the system to conceptually

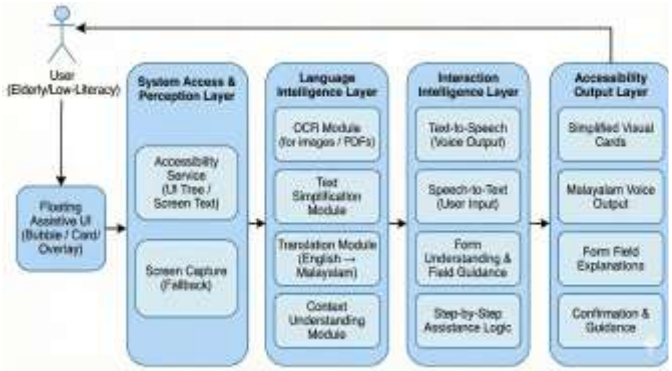

Fig. 1. Conceptual architecture of the proposed Oru Koottu digital assistant illustrating the flow of on-screen content through accessibility, language intelligence, and multimodal interaction layers.

integrate multiple AI capabilities while remaining independent of specific implementation technologies.

C. Interaction Flow

The conceptual interaction flow of *Oru Koottu* is designed to minimize user effort. When users encounter unfamiliar on-screen content, the floating interface provides immediate access to assistance. The system captures the content, processes it through the language intelligence layer, and presents simplified or translated output through visual cards or spoken feedback.

The interaction is designed to be *context-aware* rather than task-specific. Users are not required to manually copy text, switch applications, or repeatedly configure language settings. This differs from existing accessibility tools that rely heavily on explicit user actions.

D. Guided Accessibility and Safety Assistance

Beyond translation and simplification, the framework incorporates guided accessibility mechanisms to improve user confidence during digital interactions. The system conceptually utilizes accessibility APIs to identify interface components such as buttons, text fields, and navigation icons. Relevant UI elements can then be visually highlighted while voice instructions guide users through required actions.

The framework also integrates intent-aware classification and source reliability analysis to improve digital safety. Online content and extracted URLs can be categorized into semantic classes such as government notices, banking information, advertisements, or suspicious content. Malicious link detection mechanisms further warn users against unsafe or phishing-related websites.

Additionally, speech-to-text mechanisms support voice-driven form interaction by allowing users to speak information instead of typing manually. These capabilities transform the system from a translation utility into a broader accessibility assistant.

E. Support for Low-Literacy and Elderly Users

From an accessibility perspective, *Oru Koottu* emphasizes inclusive interaction mechanisms. Text-heavy explanations are replaced with concise summaries, audio narration, and step-by-step guidance. The framework also supports assistive scenarios such as form comprehension and confirmation through voice feedback.

By embedding assistance directly within the user's active screen, the system reduces dependence on reading proficiency and prior digital knowledge. This aligns with accessibility research advocating voice-based interaction and persistent guidance for elderly and low-literacy populations.

F. Distinctiveness from Existing Solutions

Unlike conventional translation applications, screen readers, or document-based assistants, *Oru Koottu* is distinguished by its *persistent, cross-application, and integrated nature*. The system does not replace existing tools but conceptually unifies their capabilities into a single assistive framework.

To the best of the authors' knowledge, existing literature does not present a unified conceptual model combining on-screen content perception, multilingual simplification, speech interaction, and accessibility guidance within a floating system-level assistant. This motivates the proposed framework and establishes its relevance within the current research landscape.

VI. Applications and Use Cases

The conceptual framework of *Oru Koottu* enables a wide range of real-world applications aimed at improving digital accessibility for elderly and low-literacy users. Unlike task-specific assistive tools, the proposed system is designed to operate across applications and usage contexts, offering continuous support wherever textual or interactional barriers arise.

A. On-Screen Content Understanding

One of the primary applications of *Oru Koottu* is assisting users in understanding unfamiliar or complex on-screen content. This includes messages, notifications, application instructions, and informational text displayed in unfamiliar languages. By capturing on-screen content and providing simplified explanations or translations, the system reduces dependence on reading proficiency and language fluency.

This use case is particularly relevant for elderly users who struggle with rapidly changing application interfaces and terminology, as well as users operating devices configured in non-native languages.

B. Multilingual Assistance for Everyday Applications

Many users interact daily with applications such as messaging platforms, browsers, banking apps, and government service portals that predominantly use English or other dominant languages. *Oru Koottu* conceptually supports real-time multilingual assistance by translating and explaining content without requiring users to leave the current application.

This cross-application translation capability addresses a key limitation of existing tools, which often require copying text or switching to separate translation interfaces, thereby increasing cognitive and operational burden.

C. Voice-Based Interaction and Information Access

Speech-based interaction forms a critical application area for the proposed system. *Oru Koottu* enables users to listen to translated or simplified content through voice output, making information accessible to users with visual impairments, reading difficulties, or age-related cognitive decline.

Additionally, voice feedback can be used to confirm user actions, read instructions aloud, or guide users through complex tasks. This aligns with accessibility research emphasizing the effectiveness of speech interfaces for low-literacy and elderly populations.

D. Form Comprehension and Assisted Data Entry

Digital forms represent a significant barrier for low-literacy users due to unfamiliar terminology, validation rules, and structured input requirements. The proposed system conceptually supports form-filling assistance by explaining field purposes, providing voice guidance, and simplifying instructions at the point of interaction.

This application is particularly relevant in contexts such as online registrations, service applications, banking forms, and government portals, where errors or misunderstandings can have serious consequences for users.

E. Support for First-Time and Non-Technical Smartphone Users

For users with limited prior exposure to smartphones, navigating modern mobile interfaces can be overwhelming. *Oru Koottu* addresses this challenge by acting as a persistent digital companion that offers contextual guidance rather than static tutorials.

By embedding assistance directly into the user's interaction flow, the system supports exploratory learning and gradual familiarization with digital services, reducing fear and dependence on external help.

F. Rural and Low-Connectivity Environments

The conceptual framework also considers deployment in rural or low-connectivity environments. By supporting offline or hybrid language processing mechanisms, the system can continue providing essential assistance even when network connectivity is intermittent or unavailable.

This makes *Oru Koottu* applicable in regions where digital inclusion is most needed but infrastructure constraints limit access to cloud-dependent services.

G. Educational and Social Inclusion Scenarios

Beyond individual usage, the system has potential applications in educational and community settings. It can support adult literacy programs, digital awareness initiatives, and community service centers by enabling users to independently access digital information.

By lowering interaction barriers, the system contributes to broader social inclusion goals, empowering users to participate more actively in digital communication, services, and information exchange.

H. Fraud Prevention and Safe Digital Interaction

The proposed framework can assist users in identifying potentially unsafe digital content such as phishing links, fraudulent websites, and misleading advertisements. By integrating source reliability verification and malicious link analysis, the system can provide warning alerts before users interact with suspicious content. This capability is particularly important for elderly and low-literacy users who are more vulnerable to online fraud and misinformation.

I. Voice-Guided Navigation Assistance

The system conceptually supports guided digital interaction through UI element interpretation and step-by-step voice instructions. Buttons, icons, and input fields can be highlighted and explained using accessibility APIs, enabling users to navigate unfamiliar applications more confidently. This feature is especially beneficial for first-time smartphone users and elderly individuals with limited digital experience.

VII. FUTURE SCOPE

While the proposed *Oru Koottu* digital assistant concept addresses several limitations identified in existing literature, multiple research and development directions remain open for future exploration.

A. Enhanced Context-Aware Understanding

Future work can focus on improving contextual understanding of on-screen content by incorporating advanced multimodal learning techniques. Integrating visual context, user interaction patterns, and semantic understanding can enable more accurate interpretation of ambiguous or fragmented text, particularly in complex user interfaces.

B. Adaptive Personalization and User Modeling

Personalization remains a critical challenge in accessibility-focused systems. Future extensions may include adaptive user modeling techniques that learn individual user preferences, literacy levels, and interaction behaviors over time. Such personalization could allow the assistant to dynamically adjust explanation depth, language complexity, and interaction modality.

C. Expanded Language and Dialect Support

Although multilingual translation models have improved significantly, future research can explore deeper support for regional dialects, code-mixed language usage, and colloquial expressions. Expanding coverage for underrepresented languages would further enhance digital inclusivity, particularly in linguistically diverse regions.

D. Improved Offline Intelligence

Future systems may explore more efficient on-device models to enhance offline capabilities. Lightweight neural architectures, model compression techniques, and edge intelligence can enable improved performance in low-connectivity environments without compromising translation or speech quality.

E. Multimodal Interaction Expansion

Beyond text and speech, future versions of the system could incorporate gesture-based interaction, visual highlighting, and haptic feedback to support users with diverse accessibility needs. Multimodal feedback can reduce cognitive load and provide more intuitive guidance during complex interactions.

F. Security, Privacy, and Trust-Aware Design

As the assistant operates at a system level, future research must address privacy-preserving mechanisms, secure data handling, and user consent management. Developing transparent interaction models and trust-aware explanations will be essential to ensure user confidence, particularly among elderly users.

G. Evaluation Through User-Centered Studies

While the current work is conceptual, future research should involve comprehensive user studies with elderly, low-literacy, and digitally marginalized populations. Longitudinal evaluations can assess usability, acceptance, learning effects, and real-world impact, contributing empirical evidence to the accessibility research domain.

H. Integration with Public Digital Services

Future extensions may explore direct integration with government portals, healthcare applications, and financial services to further reduce access barriers. Such integration could support broader societal objectives related to digital governance, healthcare accessibility, and financial inclusion.

I. Advanced Intent and Risk Prediction

Future systems may incorporate deep contextual reasoning models capable of predicting user risk scenarios, identifying financial fraud patterns, and adapting guidance dynamically based on user interaction behavior.

VIII. Conclusion

This paper presented a comprehensive literature survey on technologies and design approaches relevant to AI-driven floating accessibility and digital safety assistant for elderly and low-literacy users. The review covered key research areas including OCR-based text extraction, machine translation for low-resource languages, text simplification and summarization, speech-based interaction technologies, and accessibility-oriented user interface design.

The analysis reveals that although significant progress has been achieved within each individual domain, existing solutions largely operate in isolation. Most systems are application-specific, require explicit user intervention, or function in offline and document-centric settings. Such limitations restrict their effectiveness for users who face linguistic, cognitive, or digital literacy barriers, particularly in dynamic mobile environments.

Through comparative analysis, this survey identified a clear research gap in the absence of an integrated, system-level assistive framework capable of continuously supporting users across applications. Current approaches rarely combine real-time OCR, contextual translation, simplification, speech output, and accessibility-focused interaction within a unified and persistent interface.

Motivated by these gaps, the paper conceptually proposed *Oru Koottu*, a floating UI-based digital assistant designed to operate as an always-available accessibility layer. Rather than replacing existing applications, the system is envisioned as an assistive companion that enhances comprehension, interaction, and inclusivity by leveraging artificial intelligence in a user-centric manner.

By consolidating insights from diverse research domains, this survey establishes a structured foundation for future work in accessible mobile interaction. The findings emphasize the need for holistic, adaptive, and context-aware assistive systems that can bridge the digital divide and enable equitable access to digital services for all users.

References

[1] P. M. L. Matyila, A. Botha, R. Alberts, and G. Sibiya, “The design of accessible and usable mobile services for low literate users,” *South African Computer Journal*, 2012.

[2] M. N. Islam, N. I. Khan, T. T. Inan, and I. H. Sarker, “Designing user interfaces for illiterate and semi-literate users: A systematic review,” *SAGE Open*, 2023.

[3] S. Wickramathilaka, J. Grundy, K. Madampe, and O. Haggag, “Accessibility recommendations for designing better mobile application user interfaces for seniors,” *ACM Transactions on Computer-Human Interaction*, 2022.

[4] J. Saji, M. Chandran, M. Pillai, N. Suresh, and R. Rajan, “English-to-malayalam machine translation framework using transformers,” in *Proceedings of International Conference on Computing*, 2022.

[5] V. Atmakuria and M. Dhanalakshmi, “Improving text recognition in natural scenes using optimized ocr techniques,” in *Proceedings of ICACRS*. IEEE, 2024.

[6] O. Ferguson, “No language left behind (nllb-200) ai translation model review,” *Computer Science Review*, 2022.

[7] E. A. Emmanuel and M. H. Muyingi, “Speech user interface for low literacy users of ict services,” *European Journal of Computer Science and Information Technology*, vol. 2, no. 1, pp. 18–29, 2014.

[8] A. Jiju, S. Tuscano, and C. Badgujar, “Ocr text extraction,” *International Journal of Engineering and Management Research*, vol. 11, no. 2, 2021.

[9] J. Joseph and A. Kurian, “Breaking barriers: Transformer-based summarization and translation of english legal documents to malayalam,” in *7th International Conference on Contemporary Computing and Informatics (IC3I)*. IEEE, 2024.

[10] N. B and S. Joseph, “A hybrid approach to english to malayalam machine translation,” *International Journal of Computer Applications*, vol. 81, no. 8, 2013.

[11] R. Tom and J. Mathew, “An automated machine learning nlp model for enhancing the translation of english causative sentences into malayalam,” *International Journal of Natural Language Processing*, 2022.

[12] A. H. P., J. Kunjumon, S. R, and A. S. Ansalem, “Malayalam speech to text conversion using deep learning,” *IOSR Journal of Engineering*, vol. 11, no. 7, pp. 24–30, 2021.

[13] V. M. Reddy, T. Vaishnavi, and K. P. Kumar, “Speech-to-text and text-to-speech recognition using deep learning,” in *Proceedings of the Second International Conference on Edge Computing and Applications (ICECAA)*. IEEE, 2023.

[14] F. Ferrag and I. A. Bentounsi, “The use of artificial intelligence in academic translation tasks: Case study of chatgpt, claude and gemini,” *Journal of Linguistics and Translation Studies*, 2023.

[15] N. Reyes-Dorta, P. Caballero-Gil, and C. Rosa-Remedios, “Detection of malicious urls using machine learning,” *Wireless Networks*, vol. 30, pp. 7543–7560, 2024.