CS 528 Quick Learning App

Aikeremu Aixilafu
WPI
MSCS Student
aaixilafu@wpi.edu

Adhiraj Budukh
WPI
MSCS Student
abudukh@wpi.edu

Ester Jere
WPI
MCS Student
ejere@wpi.edu

Khushita Milind Joshi
WPI
MSCS Student
kjoshi@wpi.edu

Antonela Tamagnini
WPI
MCS Student
atamagnini@wpi.edu

ABSTRACT

The Quick Learning App distinguishes itself as a cutting-edge educational tool tailored to the unique needs of non-native English-speaking students. Setting it apart from other learning apps, Quick Learning places a strong emphasis on addressing the diverse linguistic background of its users. This innovative application empowers students to seamlessly employ their phone cameras for capturing content from various sources, including the board, presentations, and written text.

At its core, the app excels in the precise recognition and extraction of text from captured images. Additionally, it offers advanced features such as audio playback and translation into the student's preferred language. This multifaceted functionality plays a pivotal role in bridging language gaps between international students and English-speaking professors, cultivating a more inclusive and accessible learning environment.

With a thoughtfully designed user interface and efficient image processing capabilities, Quick Learning emerges as a useful companion for students dedicated to enhancing their understanding of educational materials within a multilingual context.

Categories and Subject Descriptors

H.5.2: User Interfaces
H.5.2.m: User-centered design
I.2.10: Computing Methodologies - Artificial Intelligence
I.2.10.a: Vision and Scene Understanding
K.3.2: Computers and Education - Computer and Information Science Education
D.2.11.m: Mobile architecture, Android

General Terms

The general terms of this paper include Android Development, Educational Technology, User Experience (UX) and Language Translation.

Keywords

Quick Learning App, Android App, Image Recognition, third-party libraries.

1. INTRODUCTION

In the realm of computer programming and application development, the pursuit of quick learning strategies is imperative for efficiently acquiring and applying new knowledge and skills. The effectiveness of learning methodologies is evident across various mediums, including video, text, and audio-based formats. The focus of our project, however, lies specifically in text-based features, recognizing its significance in addressing a critical challenge faced by students.

In the academic year 2022/2023, there has been a noteworthy surge in the enrollment of international students in United States universities, reflecting an increase of at least 12% compared to previous years. As of the latest available data, the United States hosts over one million international students across all academic disciplines (Open Doors, 2023).

The urgency of the issue is underscored by the increasing need to record and process classroom content in real-time. Particularly, students often find themselves in situations where rapid access to information from visual materials is crucial for reinforcing their understanding and accelerating the learning process.

The importance of solving this problem is multifaceted. Firstly, it aligns with the necessity for time-saving measures, enabling students to efficiently acquire knowledge within compressed timeframes. Additionally, these methods contribute to enhancing concentration levels, allowing students to engage more deeply with the subject matter. Furthermore, the accessibility of text-based features is a universal advantage, applicable across diverse subjects, regardless of the academic major.

This problem gains further significance in the context of our student demographic. Consequently, understanding professors during lectures becomes an arduous task, and note-taking, a traditionally relied-upon method, becomes especially difficult.

This app, designed to streamline the learning process, represents a comprehensive tool that caters to diverse learning modalities, offering students a multifaceted approach to assimilating and comprehending new information efficiently.

2. RELATED WORK

There are numerous apps categorized as quick-learning apps. During the research and development of this app, the following Android applications were employed to enhance the understanding of the principal concepts:

2.1 Google Lens

Microsoft Lens, a free app available on the Google Play Store, is a multifunctional app that converts various text formats, including images, Microsoft Office content, and handwritten notes. It allows users to save the converted text to Microsoft OneNote or OneDrive.

2.2 PenToPrint

PenToPrint is a free text application available on the Play Store, specializing in Handwriting Optical Character Recognition. It transforms scanned handwritten notes into digital text.

2.3 Microsoft Lens

Microsoft Lens, a free app on the Google Play Store, is a versatile tool for converting a range of text-based elements, including images, Microsoft Office content, and handwritten text, into digital formats. It allows for storage through integration with Microsoft OneNote or OneDrive.

3. APPROACH

The app addresses the need for instantaneous text recognition and processing in a mobile environment. It is particularly designed for users who require quick digitalization of text, whether for accessibility reasons, or for convenience in information capture and translation.

The core objective is to provide a tool that enables users to interact with physical text in a digital format seamlessly. Emphasizing user convenience, the app integrates accelerometer-based gesture controls. This decision is rooted in the principle of reducing the complexity of user interactions, enabling functions like camera activation with simple physical gestures.

Google's ML Kit was selected for its robust performance in optical character recognition. Its ability to accurately process various text formats and integrate seamlessly into the Android system, made it an ideal choice for real-time text recognition.

A critical design aspect is making the app accessible in audible format. This is achieved through the integration of Text-To-Speech technology, allowing the app to read aloud the recognized text, thus bridging the gap between written content and auditory information.

A key strategy was the fluid integration of the accelerometer sensor, camera functionality, Optical Character Recognition, and Text-To-Speech to create a cohesive and intuitive user experience. Each component complements the others, ensuring a smooth workflow from text capture to processing.

4. IMPLEMENTATION

Quick Learning, designed for efficient text recognition and processing, incorporates a blend of advanced technologies and user-friendly interfaces. Here, the key features and methodologies integrated into the app are delineated:

4.1 Accelerometer Sensor Integration

Utilizing the device's built-in accelerometer sensor, the app detects the specific motion of shaking the device, gesture leading to the activation of the camera.

4.2 Camera Operation and Control

The app employs Android's camera framework to capture images. It uses camera control features from the AndroidX Camera library, to manage camera settings and operations. This feature is central to the app's capability to capture images for text recognition (see Figure 1).

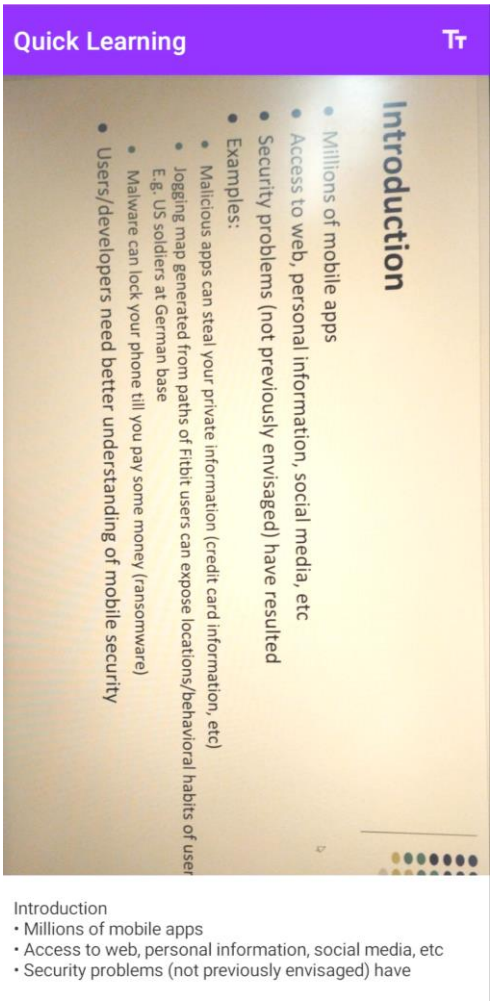

Figure 1. Main screen image capturing

4.3 Optical Character Recognition (OCR)

At the core of the app's functionality is the OCR capability, powered by Google's ML Kit. This feature enables the app to recognize and extract text from the captured images. The OCR process is optimized to handle various text formats and lighting conditions, ensuring reliable text extraction (see Figure 2).

4.4 Text-to-Speech (TTS) Integration

Post text recognition, the app employs Android's Text-to-Speech engine to read aloud the recognized text. This feature enhances accessibility, particularly benefiting users with visual impairments or those seeking a hands-free experience.

4.5 Translator Feature

Recognizing the diverse linguistic needs of users, the app includes a feature to translate the recognized text. It creates an intent to open a web-based translation service, allowing users to easily obtain translations of the extracted text.

4.6 Sharing Functionality

To facilitate easy documentation and sharing of the output, the app integrates a feature to send the text to other multiple apps. This is achieved through a share intent, which makes the text

available to other apps installed on the device, including various note-taking platforms (see Figure 3).

Figure 2. Optical Character Recognition functionality

5. EVALUATION

5.1 Efficient Text Recognition and Extraction

Utilizing Google's ML Kit for Optical Character Recognition, the app demonstrates high proficiency in accurately recognizing and extracting text from images. This feature is central to the app's functionality and works as expected, catering to the needs of students who require quick digitalization of materials.

5.2 Audio Playback through Text-to-Speech

The integration of Text-to-Speech technology effectively converts recognized text into audio format. This not only aids in better comprehension for auditory learners but also makes the app accessible to visually impaired users. The functionality appears to work seamlessly, providing a clear and understandable audio playback of the text.

5.3 Multilingual Support and Translation

The app's ability to access Google Translator through a button to translate the recognized text into various languages addresses the diverse linguistic backgrounds of its users. The decision to use a web-based translation service for this feature ensures broad language coverage and provides reliable translation quality.

5.4 User-Friendly Design with Accelerometer-Based Gesture Controls

The app's design, which includes gesture controls using the device's accelerometer, enhances the user experience by simplifying interactions. This feature works as expected, allowing users to activate the camera with intuitive physical gestures.

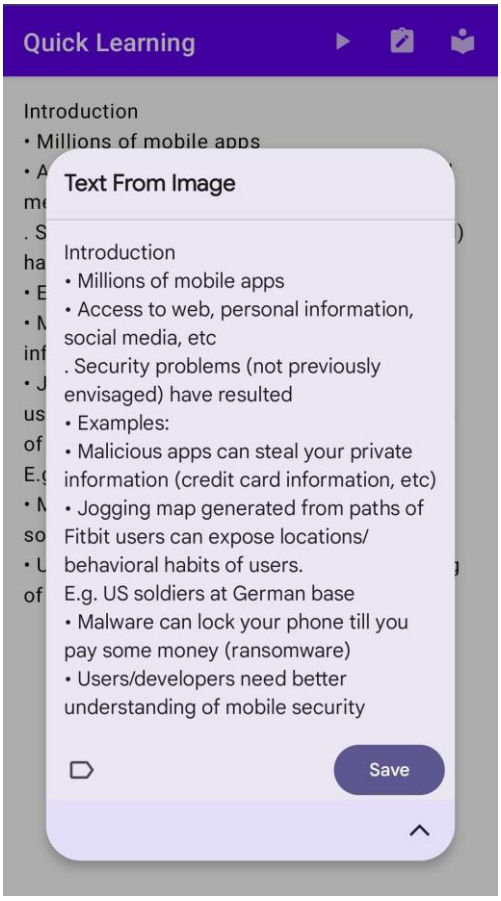

Figure 3. Share functionality: note-taking feature

5.5 Seamless Integration of Camera for Text Capture

The use of Android's camera framework and the AndroidX Camera library for image capture ensures a robust and reliable operation. This aspect of the app is critical for capturing clear images for text recognition.

5.6 Note-Taking and Sharing Functionality

The app's feature to share recognized text with note-taking and other applications adds considerable value, especially for students who rely on digital notes. This functionality facilitates easy documentation and organization of lecture materials.

5.7 Overall Cohesive and Intuitive User Experience

The integration of various technologies - accelerometer sensor, camera functionality, OCR, TTS - stands fluid and cohesive, creating an intuitive user experience. Each component complements the others, ensuring a smooth workflow from text capture to processing and translation.

6. DISCUSSION

In the development of the Quick Learning, the integration of camera functionality has been noted to offer an innovative approach to learning through visual content. The mechanism of

shaking the phone to activate the camera provides a user-friendly and intuitive way for students to engage with the application.

The incorporation of audio playback and sharing options facilitate a multi-sensory learning experience, catering to different learning preferences. It enables users to manage the content on the image, enhancing comprehension and accessibility.

An open problem of the application is the inability to recognize handwritten text for all types of fonts. Currently, the app detects and recognizes only clear and legible handwriting. Initially, to address this issue, we considered using ML Kit's Digital Ink Recognition feature to support all handwritten text recognition, but the core technology and design of Digital Ink Recognition are fundamentally incompatible with processing static images of text captured by a camera. This ML Kit feature is specifically designed to recognize handwritten input on a digital surface, such as a touchscreen, where the user physically writes or draws using their finger or a stylus. This process involves analyzing the strokes made by the user in real-time or from recorded touch inputs. In this regard, the stroke data in digital ink includes temporal information (like the duration and sequence of each stroke), which is critical in understanding handwriting. This information is absent in a static image captured by a camera. Furthermore, Digital Ink Recognition can include context about the writing area and expected text orientation based on how the user interacts with the touchscreen. Camera-captured images do not inherently provide this contextual information, making it difficult for the algorithm to accurately interpret the text.

7. CONCLUSION

The Quick Learning app's innovative approach to integrating advanced technologies like OCR, TTS, and translation services within a user-friendly mobile application, presents a powerful tool for improving educational accessibility and efficiency. It stands out as a valuable resource for international students and those who face language barriers, aligning with the increasing diversity in global education settings.

8. FUTURE WORK

Future development of Quick Learning should focus on refining the flexibility of input, not only limiting it to text but also including images, videos, and audio. Additionally, incorporating Artificial Intelligence would enable the app to search for related items on the web based on the input provided by the user. Another consideration is the inclusion of collaboration features to facilitate group work, allowing students to view the same data simultaneously and contribute collectively. Integrating cloud services, such as Amazon Web Services (AWS), could enhance data storage and sharing capabilities, enabling users to access their learning materials across multiple devices.

Furthermore, expanding platform support to include iOS and other Android versions would broaden the app's accessibility.

9. REFERENCES

[1] Damyanov, I., & Tsankov, N. (n.d.). Mobile Apps in Daily Learning Activities. *International Journal of Interactive Mobile Technologies*, 12(6), Article 9659. https://doi.org/10.3991/ijim.v12i6.9659

[2] Jackson, M., Ray, S., & Bybell, D. (2013). International students in the U.S.: Social and psychological adjustment. *Journal of International Students*, 3(1), 17-28.

[3] Open Doors. (2023). International students in the United States: Total enrollment trends. *Retrieved from* https://opendoorsdata.org/data/international-students/

[4] Almurideef, Raja. (2016). The challenges that international students face when integrating into higher education in the United States. *Theses and Dissertations*. 2336. https://rdw.rowan.edu/etd/2336

[5] Wu, H.-p., Garza, E., & Guzman, N. (2015). International Student's Challenge and Adjustment to College. *Education Research International*, 202753. https://doi.org/10.1155/2015/202753

[6] B. Sekar and J. B. Liu. (2014). Location based mobile apps development on Android platform. *9th IEEE Conference on Industrial Electronics and Applications, Hangzhou, China*, pp. 2148-2153, doi: 10.1109/ICIEA.2014.6931527. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6931527&isnumber=6931119