Enhancing Document Digitization: The All-in-One 'Document World' App for Scanning and Editing PDFs with QR Code Support

Utkal University
SSPMâs College of Engineering, Kankavali.
Sindhudurg, Maharashtra, India, 416602.
[1]tanmay.kadam@SSPMCollege.ka.in

Abstract. This article aims to support students, employees, and the general public by leveraging document scanning, optical character recognition (OCR), QR code generation, and QR code reading. With the growing prevalence of digitalization in educational institutions, offices, and governments, there is a consistent demand for efficient tools to fulfill these tasks. To address this need, we advocate for a dependable scanning application equipped with essential features like PDF manipulation tools, QR code generation, and QR code reading capabilities. While numerous online applications exist, they often fall short in terms of quality and functionality, sometimes necessitating payment for advanced features. However, the "Documents World App" stands out as a viable solution, offering high-quality scanning, PDF tools for modifying existing documents, and editing features for adding signatures and watermarks, streamlining these processes. Noteworthy are its advanced functionalities, including text extraction facilitated by the itext library, a feature not commonly found in widely used tools such as "Adobe Scan" and "iLove PDF." Crucially, our objective is to provide this service to all users without any charges and intrusive advertisements, ensuring accessibility and affordability for everyone.

Keywords: Document Scanning, PDF Document Management, QR, OCR, PDF tool, Mobile App.

1 Introduction

Numerous prior studies have indicated a continuous rise in the percentage of the population utilizing mobile technologies. This upward trend is notably fueled by the widespread adoption of the Android operating system and the increased accessibility to the internet. Mobile phones have evolved into essential devices in daily life, extending to various business activities. The growing reliance on mobile phones as gadgets for storing diverse documents underscores the importance of addressing document management applications on the Android platform. In contemporary times, mobile applications serve as the primary medium for facilitating human interaction. Given their widespread usage, it is essential to ensure that these applications are user-friendly for everyone. Accessibility guidelines provide developers with recommendations to incorporate features that enhance the interaction experience for users with disabilities, such as those with visual impairments [13]. The documentation process constitutes a vital aspect of daily operations in workplaces, households, and various industries. However, the simultaneous management of extensive documentation poses significant challenges, consuming excessive time and posing environmental hazards. Consequently, addressing these issues becomes imperative. While the contemporary scanner, as we recognize it today, was introduced over 50 years ago, the concept itself dates back much further. In the 1860s, the invention of the Pan-telegraph allowed individuals to transmit handwritten notes, drawings, and signatures through telegraph wires, illustrating the historical roots of document scanning. It was

frequently utilized as a means of verifying signatures in financial transactions. However, as the demands of the industry surpassed the capabilities of the Belingrophe, the development of the flatbed scanner became imperative. These scanners, which are now ubiquitous, gained popularity in the early 1990s, as previously discussed. Flatbed scanners employ optical scanning to convert handwritten documents or photographs into digital files, a functionality widely utilized by businesses worldwide. In the contemporary landscape, scanning plays a pivotal role in daily operations, facilitating the organization and retrieval of relevant data without the need for extensive storage facilities. Notably, an Android app with distinct features tailored to the needs of both students and employees has been evolved [8]. Numerous employees and students sought an application capable of scanning documents and converting them into various file formats. This app goes beyond basic scanning, offering additional features such as OCR, QR scanning, and barcode scanning [2]. As a result, our application proves highly beneficial to a diverse user base, encompassing students and employers alike. Given the prevalent shift towards online learning and remote work, particularly amidst the ongoing COVID-19 pandemic [14], the role of digital tools, particularly scanners, has become increasingly vital. While scanners have become commonplace in workplaces, educational institutions, and universities, certain critical elements are often lacking in existing scanning apps, such as high-quality scanning. Our application addresses these gaps by consolidating essential capabilities, including scanning, file format conversion, OCR, and more, into a single, user-friendly platform. This holistic approach offers users a comprehensive solution to their documentation needs. In light of the escalating demand for online work, various industries, including business and education, recognize the significance of digitizing existing systems. Upgrading the current system and enhancing its quality emerges as the most straightforward solution to the challenges posed by the increasing volume of online work. Our application not only facilitates this transition but also acknowledges the necessity of converting traditional paper documents into a digital format. With features like a QR reader and QR scanning, our application empowers users across diverse sectors to capture high-quality document images, providing a centralized hub for editing and file conversion options. Notably, this application is compatible with all Android smartphones and tablets. Mobile applications, commonly known as apps, are extensively employed for both social and emergency connectivity. The continuous surge in their usage is attributed to the growing practice of

social distancing. These apps are pivotal in facilitating human interaction, prompting researchers to focus on creating tools that aid in the development of effective mobile applications accessible to diverse user groups. This fact motivated us to work in this research domain. Hence, the research presented in this article endeavors to tackle pertinent issues, aiming to achieve the following key highlights:

Highlights

- Centralize all essential digital document features within the application.

- Streamline the management of PDF documents within the application.

- Improve the scanning quality of PDF documents.

- Design a user-friendly interface for the App to enhance its usability.

- Introduce a free version of the app.

The article comprehensively addresses the state-of-the-art literature review in the specific research domain, as well as delves into the methodology, implementation, results, and discussion in the subsequent sections.

2 Related work

The literature concerning the selected research domain is indeed limited in the available publications. Consequently, our search endeavors were concentrated on identifying and extracting relevant information within this research domain. A thorough literature review was undertaken to pinpoint opportunities for improvement within this domain. Our examination of existing publications provided valuable insights into the complexities of Android app development [5], with a specific focus on the use of an SQLite database [7] and OCR [3] in scanning applications. The knowledge gleaned from this literature review will be incorporated into our project to augment its overall functionality. The research presented in [4], aims to delve into the intricacies of document management on Android devices. The research presented in [1], centers on the implementation of a document management system at Romblon State University, Cajidiocan Campus, located in Romblon, Philippines. Currently, educational institutions, including Romblon State University, generate a substantial volume of documents such as feedback reports, meeting minutes, syllabi, lesson plans, policies, memos, and circulars. Traditionally, these documents are stored in

filing cabinets and serve as compiled evidence for accomplishment reports or accreditation purposes. However, this storage method presents challenges related to space, document security, and the efficient sharing of documents when required. Consequently, this study aims to address these issues by facilitating easy access to documents using social technology across the academic organization and by providing expanded storage capabilities through cloud computing. The research article [12] introduces a novel image segmentation algorithm employing an adaptive split-and-merge approach grounded in characteristic features and a hypothesis model. The Android app proposed in [10] facilitates file sharing between different systems, rendering it a distinctive Android project. The document scanning aspect relies on the Canny edge detection algorithm, while the file-sharing functionality operates through WiFi Direct, establishing a server-client connection to facilitate seamless file transfers. Despite forgoing certain extra features present in alternative scanning systems, this project prioritizes the accurate creation of documents and the uninterrupted, secure transfer of files between systems. It is important to acknowledge potential limitations, such as the speed of the transfer action. The utilization of OpenCV plays a pivotal role in edge detection for document imaging perspective transformation. This transformative process includes image rectification through a four-point transformation method, culminating in the generation of high-quality processed images suitable for compilation into a PDF file. An integral aspect of the research presented in [6] involves the proposition of a fully automatic preprocessing technique designed to elevate the quality of digitally captured images by rectifying various types of distortion introduced by digital cameras. The primary objective of [2] is to conduct a comprehensive examination of the SQLite Android relational database management system, assessing its performance across various scenarios. To facilitate this investigation, an application exemplifying performance dynamics within the SQLite database has been meticulously developed and made available on the Google Play Store. The empirical findings derived from the testing process reveal a notable discrepancy in the duration of operations between encrypted and unencrypted data, indicating that operations on encrypted data incur significantly longer processing times. OCR enjoys popularity within Android applications, and Tesseract stands out as a widely utilized open-source library for the implementation of OCR in Android applications [11]. Chandni Kaundilya et al. [?] focus on the extraction of text from images through a series of techniques,

including text localization, segmentation, and binarization. The process of text extraction begins with text detection, which identifies image segments containing text. Following this, text localization precisely determines the position of the text within the identified segments. Subsequently, text segmentation is employed to separate the text from its background, and a binarization process is applied to convert colored images into binary format. The final step involves character recognition applied to the binary image, converting it into ASCII text. Jamshed Memon et al. [9] Systematic Literature Review (SLR) systematically gathered, synthesized, and analyzed research articles about handwritten OCR and closely related topics published between 2000 and 2019. Utilizing widely recognized electronic databases and adhering to a pre-defined review protocol, we employed keyword searches, forward reference searches, and backward reference searches to comprehensively identify relevant articles. Following a meticulous study selection process, 176 articles were included in this SLR. The objective of this review article is to present state-of-the-art results and techniques in OCR while also delineating research directions by shedding light on existing research gaps. The research presented in [2] examines the long-term evolution of mobile app usage. Through this platform, authors have compiled app usage records spanning from 2012 to 2017, encompassing data from 1,465 users.

Hence, we conducted an extensive literature review in this domain and, to the best of our knowledge, identified limited existing literature in this research area. Motivated by this research gap, we decided to explore and contribute to this domain, leading to the development of our Android app.

3 Methodologies

In this segment, we outline the methodologies employed in the creation of the Android application titled "Document World," designed to manage diverse document-related operations. The methods we propose accept input in the form of text, images, and handwritten documents. Our Android app processes this input data and produces enhanced outputs, as outlined below and illustrated in Figure 1.

- Scanned PDF document

- QR Code Generator

- Extract text from the document

- Extract images
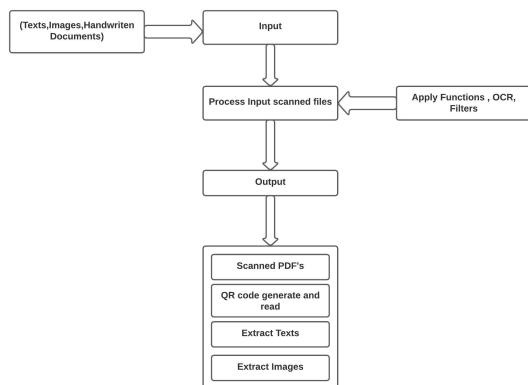
Figure 1: Proposed system's workflow

The app's implementation leverages several libraries sourced from Android Studio, including Scan Library, Smart Cropper, Camera View, Image Picker, PDF Viewer, Vision, iText, and File Storage. These libraries are described in brief in the following subsections.

3.1 Scan Library

The implementation of the Document Scanner relies on the utilization of the Scan Library, an Android document scanner library built upon OpenCV. This library enables users to select precise edges, crop documents based on user-defined four corners, and adjust the transformation of the cropped image from a perspective with the assistance of the application.

3.2 Smart Cropper

Smart Cropper, an intelligent library, can identify the borders of a cropped image and make necessary adjustments. It is well-suited for applications involving identity cards, business cards, paperwork, and other types of cropped photos.

3.3 Camera View

The open-source camera view is extensively documented and features a high-level library. It simplifies the process of capturing pictures and videos. Notably, Camera View is endorsed and supported by Share Chat.

The camera engine will notify you about camera events that took place. We can access these events using the *CameraListener* instance[?]. We can add listeners to the *CameraView* using *addCameraListener()*. One of the events we are interested in is *onPictureTaken()*. When the image clicked, this event occurred.

3.4 Image Picker

The Android library facilitates image selection from the device through a customizable user interface, applicable to both activities and fragments. It supports the selection of various image formats, such as JPEG, PNG, GIF, and video formats, including MPEG and MP4 [?]. Users can apply different themes, choosing from two built-in themes or creating custom themes. This library serves the purpose of enabling users to select an image from the device, which can be subsequently utilized for creating scanned images or PDF documents.

3.5 PDF Viewer

An Android library designed for the presentation of PDF documents, featuring animations, gestures, zoom capabilities, and support for double-tap interactions [?].

3.6 Vision

The text recognition API [?], is capable of recognizing text in any Latin-based character set. Key capabilities include:

- Recognize text across Latin-based languages

- Analyse the structure of a text.

- Identify the language of the text.

- Real-Time Recognition

3.7 iText

iText is a Java-based PDF library that allows the development of Java programs for creating, converting, and manipulating PDF documents [?]. It offers classes (APIs) for generating interactive PDF documents, enabling the creation of maps and books. Additionally, iText supports operations such as splitting an existing PDF into multiple documents and adding or concatenating additional pages. As iText operates as a free and open-source software (F/OSS) initiative, its source code is accessible without charge and is governed by the AGPLv3 license. Within the AGPLv3 framework, you are permitted to utilize the iText Core (alternatively referred to as iText Community) PDF library and our open-source add-ons without incurring any expenses.

3.8 File Storage

Android employs a file system that closely resembles disk-based file systems found on other platforms. The file storage mechanism provides a means to store data

persistently on the user's device[?] [?]. Figure 2 illustrates the visual depiction of the process by which data storage occurs

Figure 2: Illustration of file storage mechanism

4 Implementations

The features within the Documents World Android application fall into the following five categories:

1. Document Scanner

2. PDF Tools

3. Image to Text, i.e. OCR

4. QR Generator

5. QR Reader or Scanner

A comprehensive explanation of the implementation of these categories is described in this section. Users can scan, store, and edit documents. Additionally, advanced features such as custom color filters, document signing, text addition, highlighting, and more are also accessible to users. Various PDF tools designed to improve user convenience are available [?]. These tools take PDF or text files as input, perform specified operations, and provide the desired outputs[?] [?]. Some important tools are outlined below:

1. Merge and Split PDFs

2. Remove and Reorder Pages

3. Extract Images

The algorithm processes an image provided by the user and extracts the text within it, as demonstrated in Figure 3. Regardless of the presence of design elements or other graphics in the image, the algorithm successfully extracts text in any font [?], [?].

Figure 3: Image to Text

We have developed a straightforward QR Code generator app where users can input data within the app, and the generated QR Code will display the incorporated data. Users can generate QR (Figure 4) of Text, E-mail, Phone, SMS, or any URL. Users will also be able
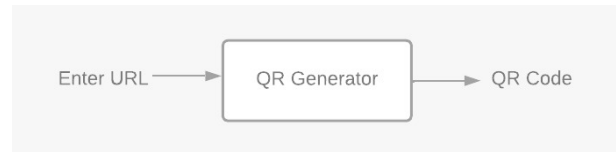
Figure 4: URL or Text to QR

to scan an already existing QR to view the exact information. We have applied the coding to handle the iText Library, incorporating images into PDFs, and managing the addition and removal of passwords using Android.

We have placed a greater emphasis on improving the quality of the scanned document by adjusting the ColorMatrix, Contrast, and Exposure Matrices. The functionality and application of these matrices in the development of our proposed model are elaborated on in the subsequent subsections.

4.1 ColorMatrix

A 4 X 5 matrix is utilized to transform the color and alpha components of a Bitmap. This matrix can be provided as a single array and is interpreted as follows:

$$[a, b, c, d, e,$$
$$f, g, h, i, j,$$
$$k, l, m, n, o,$$
$$p, q, r, s, t]$$

When utilized on a color [R, G, B, A], the resulting color is calculated as:

$$R^{'} = a*R + b*G + c*B + d*A + e; \tag{1}$$

$$G^{'} = f*R + g*G + h*B + i*A + j; \tag{2}$$

$$B^{'} = k*R + l*G + m*B + n*A + o; \tag{3}$$

$$A^{'} = p*R + q*G + r*B + s*A + t; \tag{4}$$

4.2 Contrast

Contrast pertains to the distinction between the darkest and brightest areas in a photo. Elevating the contrast intensifies shadow darkness and brightens highlights. Conversely, reducing contrast brings shadows

and highlights closer together. The contrast matrix is represented as follows:

$$Contrast = ccccccc\,RGBAW\,c(ccccc)RC000tG0C00tB00C0tA00010$$

where C is the contrast and the t equals $(1.0 - c/2.0)$

4.3 Exposure

Exposure is highly biased towards the highlights in the picture. It will act as if increasing the amount of light in the picture, without much affecting the darker areas in the picture. The Color matrix to represent the exposure is shown below:

$$Exposure = ccccccc\,RGBAW\,c(ccccc)R20000G02000B$$

5 Results and Discussion

In this section, we introduce the functionality and associated outcomes of our developed Android app for document scanning. The appearance of the app's landing page and main page is illustrated below in Figure 5 5 (a). Upon opening the app, users will be greeted with this page, which showcases icons representing various features on the main screen, as shown in Figure 5 (b). Upon selecting the camera icon, as depicted in Fig-

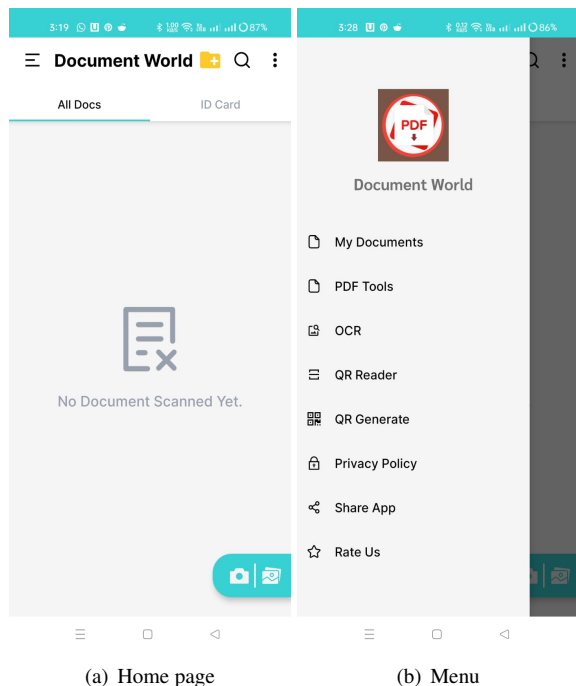
(a) Home page   (b) Menu

Figure 5: Front End of Document World Application

ure 6 (a), users will be directed to a page where they

can toggle the flashlight on or off using the corresponding icon. The flashlight is initially turned off. Positioned at the bottom right corner, a checkmark button enables users to save the scanned image and conclude the scanning process. Subsequently, by selecting the button shown in Figure 6 (b), users have the option to establish a custom folder and add scanned documents to it. Consequently, a new folder will be created within the Document Scanner folder in the mobile files. Other scanned documents can be relocated to this newly created folder using the move option. The layout menu, accessible by clicking on the keytab menu option, provides users with additional functionalities. Upon select-
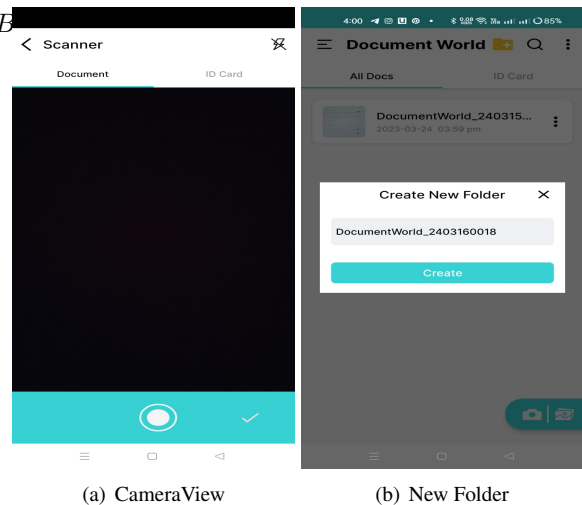

(a) CameraView   (b) New Folder

Figure 6: Screenshot of Document World App's Features

ing the hamburger button, illustrated in Figure 7 (a), users gain access to the main features of the application. Clicking on the menu button, as shown in Figure 7 (a), opens a menu displaying all the primary features. By opting for the ID Card feature in Figure 7 (b), users can generate scanned copies of ID cards. Recognizing that users often scan various types of ID cards and prefer creating custom folders for storage, a dedicated section for ID cards has been established. Since ID cards can be either one-sided or two-sided, the application prompts users to specify whether they are scanning a one-sided or two-sided ID card during the scanning process. The Image to Text feature enables users to extract and copy text from images, even when the text is typically non-editable. These features are shown in Figure 8 (a). Regardless of the font used in the image, this feature allows users to directly copy text, saving time that would otherwise be spent manually typing. This not only streamlines tasks but also allows users to concentrate on more critical activities. The extracted text can

(a) Main features Menu   (b) ID Card
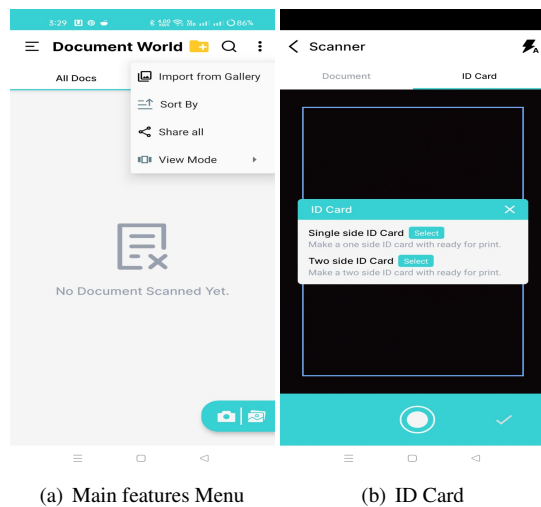
Figure 7: Depiction of app's main features: menu and ID card.

be easily shared or copied for further use. The QR code

(a) Image to Text   (b) QR Code Generator
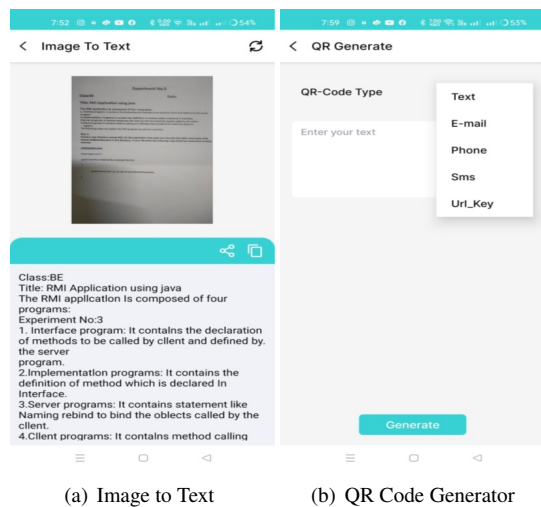
Figure 8: Depiction of app's main features menu for the image to text and QR code generation.

generator view of the app is shown in Figure 8 (b). This feature empowers users to create their own QR codes for various applications. Users can generate QR codes with:

- Text
- E-mail
- Phone Number
- SMS

- URL_Key

At times, certain passwords, such as those for Wi-Fi access, entry passes, or attendance marking, cannot be directly provided to users. Authenticators may prefer to keep these passwords specific to an area or group. In such cases, an authenticator can convert a password into a QR code and distribute it to users, ensuring the password remains concealed. Presently, users often need to generate QR codes for UPI payments. Utilizing the URL_Key option allows users to create a QR code for a URL linked to UPI payments.

6 Performance Analysis

We performed a performance evaluation comparing our implemented app with other established applications, including Adobe Scan, Doc Scanner, and iLovePDF. Our assessment uncovered the absence of certain features, such as text extraction, in these existing apps. Table 1 provides a comparative overview of the functional features of Adobe Scan, Doc Scanner, iLovePDF, and the "Document World App." In Table 1, the ✓denotes the availability of a feature in an application, while the ✗indicates the absence of that feature in the respective application. Upon evaluating the performance of the implemented Android app named "Document World" in comparison to existing apps such as iLovePDF, Doc Scanner, and Adobe Scan, it was observed that "Document World" outperforms the existing apps in certain aspects. Particularly, it excels in features such as adding or extracting text, QR reader, and generator functionalities.

7 Conclusion and Future Scope

In the rapidly evolving digital landscape, effective management of digital documents is paramount. This research project introduces the "Document World" Android app, addressing the need for efficient handling of digital documents through essential services like merging, splitting, editing, and scanning. Additionally, it offers advanced features such as OCR, QR generation, and text-to-speech capabilities. The comprehensive functionality of the "Document World" app meets various requirements, simplifying tasks such as scanning pages, modifying existing PDFs with versatile PDF tools, enhancing PDFs, and incorporating editing options like watermark addition. The integration of libraries such as Scan Library, Smart Cropper, Camera View, Image Picker, Vision, and iText has empowered the app with diverse capabilities. A comparative analysis against popular apps like Adobe Scan and iLovePDF demonstrates that the "Document World" app outshines

Table 1: Comparing Document World with Other Document Scanning Applications

Features	Adobe Scan	Doc Scanner	iLovePDF	Proposed App
Image to PDF	✓	✓	✓	✓
Text to PDF	✗	✓	✗	✓
Add password	✓	✓	✗	✓
Remove password	✗	✓	✗	✓
Add text	✗	✗	✗	✓
Rotate pages	✓	✓	✓	✓
Add Watermark	✗	✓	✓	✓
Add Images	✓	✓	✗	✓
Merge PDF	✓	✓	✓	✓
Split PDF	✗	✓	✓	✓
Remove Pages	✓	✓	✗	✓
Reorder pages	✓	✓	✗	✓
PDF to Images	✗	✓	✓	✓
Extract text	✗	✗	✗	✓
OCR	✓	✓	✓	✓
QR Reader	✗	✓	✗	✓
QR Generate	✗	✓	✗	✓
Free of cost	✗	✗	✗	✓

with its broader feature set, addressing shortcomings in current scanning and file-conversion systems. Notably, its compatibility with all Android devices positions it as a valuable tool for a wide range of users, including students and professionals. Looking ahead, the inclusion of text translation in multiple languages is envisioned to enhance the app's utility. Additionally, a planned feature for extracting handwritten text from documents and converting it into digital text promises to further elevate the app's functionality. In summary, the "Document World" app emerges as a comprehensive solution that not only meets current digital document needs but also positions itself for continued relevance and expansion in the dynamic digital ecosystem.

The suggested approach can be enhanced by incorporating features such as improving text recognition in OCR. Elevating the resolution of scanned images will be part of the user interface in the implemented version of this app and is identified as a potential avenue for future research in this work.

References

[1] Bacuna, R. D. and Castro, E. T. Acceptance of google classroom technology among students at romblon state university-cajidiocan campus using the modified utaut model. In *2023 3rd Asian Conference on Innovation in Technology (ASIAN-CON)*, pages 1–8. IEEE, 2023.

[2] Basker, E. Raising the barcode scanner: Technology and productivity in the retail sector. *American Economic Journal: Applied Economics*, 4(3):1–27, 2012.

[3] Biten, A. F., Tito, R., Gomez, L., Valveny, E., and Karatzas, D. Ocr-idl: Ocr annotations for industry document library dataset. In *European Conference on Computer Vision*, pages 241–252. Springer, 2022.

[4] Daraghmi, E., Qaroush, Z., Hamdi, M., and Cheikhrouhou, O. Forensic operations for recognizing sqlite content (forc): An automated forensic tool for efficient sqlite evidence extraction on android devices. *Applied Sciences*, 13(19):10736, 2023.

[5] Dos Santos, M. R., Rodriguez, D. Z., and Rosa, R. L. A novel qoe indicator for mobile networks based on twitter opinion ranking. In *2023 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 1–6. IEEE, 2023.

[6] Dumka, A., Ashok, A., Verma, P., and Verma, P. *Advanced digital image processing and its applications in Big Data*. CRC Press, 2020.

[7] Fayyad-Kazan, H., Kassem-Moussa, S., Hejase, H. J., and Hejase, A. J. Forensic analysis of whatsapp sqlite databases on the unrooted android phones. *HighTech and Innovation Journal*, 3(2):175–195, 2022.

[8] Goumopoulos, C., Nicopolitidis, P., Gavalas, D., and Kameas, A. A distance learning curriculum on pervasive computing. *International Journal of Continuing Engineering Education and Life Long Learning*, 27(1-2):122–146, 2017.

[9] Memon, J., Sami, M., Khan, R. A., and Uddin, M. Handwritten optical character recognition (ocr): A comprehensive systematic literature review (slr). *IEEE Access*, 8:142642–142668, 2020.

[10] Montagnetti, R. and Guarino, G. From qgis to qfield and vice versa: How the new android application is facilitating the work of the archaeologist in the field. *Environmental Sciences Proceedings*, 10(1):6, 2021.

[11] Selvakanmani, S., Chandrashekar, T., Federick, N. D., and Jaffar, A. M. Optical character recognition based text analyser: A case study. *SCIENCE*, 2020.

[12] Srivastava, V. and Biswas, B. Lm-mfp: large-scale morphology and multi-criteria-based feature pooling for image parsing. *Soft Computing*, 26(13):6201–6218, 2022.

[13] Teodoro, A. A., Silva, D. H., Rosa, R. L., Saadi, M., Wuttisittikulkij, L., Mumtaz, R. A., and Rodriguez, D. Z. A skin cancer classification approach using gan and roi-based attention mechanism. *Journal of Signal Processing Systems*, 95(2-3):211–224, 2023.

[14] Teodoro, A. A., Silva, D. H., Saadi, M., Okey, O. D., Rosa, R. L., Otaibi, S. A., and Rodríguez, D. Z. An analysis of image features extracted by cnns to design classification models for covid-19 and non-covid-19. *Journal of signal processing systems*, pages 1–13, 2021.