iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.70058

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Seamless Cross-Platform Document Conversion Application

Dr. S. S. Shriramwar[1], Chetak Chandewar[2], Kshitij Bhawalkar[3], Pratham Selokar[4], Rahul Bhoyar[5]
[1, 2, 3, 4, 5]*Dept. of Industrial IoT, Priyadarshini College of Engineering, Nagpur*

Abstract: ***The need for a versatile mobile and web application that integrates document and multimedia conversion, and accessibility features has increased due to advancements of digital technologies. This research presents the design and development of a cross-platform mobile application built using Flutter, utilizing cutting-edge technologies to meet various users' requirements. The application includes a set of tools such as PDF-to-Word, Word-to-PDF, and PPT-to-PDF converters, images to PDF, XLS-to-PDF, merge PDFs, image-based text extraction through Optical Character Recognition (OCR), speech-to-text, and video summarization through YouTube link parsing. The application also offers access to several AI tools needed for different tasks through a single search interface, enabling efficient discovery and use of AI-based features by users. This app has a scalable architecture that utilizes Flutter's responsive design capabilities to ensure optimal usability on browser, tablet, and mobile devices. The Provider state management solution provides efficient state management for seamless navigation and interaction. Integrating AI-driven capabilities, such as summarization and OCR, enriches user experience by offering precision and automation in data processing. This paper outlines the technical realization, architectural design, and development issues.***
Keywords: Document Conversion, OCR, PDF, PPT, Excel, Summarization, AI Tools, Image to PDF, PDF to Word

I. INTRODUCTION

Managing digital documents efficiently has become a necessity in both professional and personal settings. People often work with multiple file formats like PDFs, DOCX, and images containing text. However, converting files and extracting information usually requires different tools, making the process time-consuming and inefficient. To simplify this, FileSmith was developed as an all-in-one document management app, combining speech-to-text conversion, text extraction from images, and PDF to DOC conversion into a single, easy-to-use platform. It also has features such as images to PDF conversion, merging multiple PDFs, PPT to PDF conversion, PDF to PPT conversion, and DOC to PDF conversion. One of the highlighting features of the application is that it provides access to various AI tools needed for different tasks through a single search system, making it easier for users to find and use AI-powered features efficiently. The speech-to-text feature allows users to convert spoken words into editable text using advanced speech recognition. This is especially useful for note-taking, transcription, and accessibility support. Meanwhile, the image-to-text extraction tool uses Google ML Kit to recognize and extract text from images, helping users quickly digitize printed or handwritten content. The extracted text can be edited, copied, or saved as a PDF for future reference. Another essential feature is the PDF to DOC converter, which provides a simple way to transform PDFs into editable Word documents. This feature is integrated with an online processing server, ensuring faster and more accurate conversions without overloading the device's processing power. Built using Flutter and Firebase, FileSmith ensures seamless performance across multiple platforms while allowing users to control where their files are stored, optimizing space and accessibility. By bringing together essential document management features in a single app, FileSmith eliminates the hassle of using multiple software solutions, making it an efficient tool for students, professionals, and everyday users looking for a smarter way to handle digital documents.

II. LITERATURE REVIEW

The demand for digital tools that facilitate efficient document management and conversion has been a recurring theme in technological advancements. Existing solutions often address isolated functionalities, such as file format conversion or OCR, but fail to offer a cohesive platform that integrates multiple capabilities. This section reviews prior work and identifies the gaps this research aims to fill.

A. Document Conversion Tools

Applications like Adobe Acrobat and Small PDF have long dominated the document conversion space, providing reliable tools for converting files between PDF, Word, Excel, PowerPoint, and other formats.

However, these solutions are often platform-specific or require expensive subscriptions, limiting their accessibility for broader audiences. Additionally, standalone tools for merging PDFs or handling specialized conversions are often fragmented. The proposed application addresses these limitations by offering a free, cross-platform solution that integrates various document conversion capabilities, including merging PDFs.

B. Optical Character Recognition (OCR)

OCR technology has seen significant advancements, with tools like Google's Tesseract achieving impressive accuracy levels. Despite these advancements, OCR tools often operate as standalone applications, lacking integration with other document processing functionalities. The integration of OCR within a versatile document conversion platform bridges this gap, enabling users to extract and utilize text seamlessly from images, enhancing productivity in academic, professional, and personal contexts.

C. Speech-to-Text Technology

Accurate transcription is possible using speech-to-text systems, as those provided by IBM Watson and Google Cloud Speech-to-Text. While effective, these solutions often rely on robust internet connectivity and cloud-based processing, which can be resource-intensive. The proposed application integrates this functionality to convert spoken content into editable text while optimizing performance for use on mobile and low-resource devices.

D. Summarization Tools

Text summarization has gained traction with the advent of transformer-based models like BERT and GPT. Current implementations, however, are often confined to research settings or standalone APIs. Integrating AI-driven summarization within a document management app not only enhances functionality but also aligns with the growing need for automated content processing. Summarization can assist researchers, students, and professionals in extracting essential insights from lengthy documents and presentations.

E. Cross-Platform Development

The emergence of frameworks like Flutter and React Native has revolutionized cross-platform app development. Flutter's capability to deliver consistent performance across Android, iOS, and web platforms makes it an ideal choice for this project. Unlike traditional native development, Flutter reduces development time and ensures feature parity across platforms. This enables the proposed application to cater to users on diverse devices, ensuring broad accessibility.

F. Gaps Identified

While individual tools exist for document conversion, OCR, speech-to-text, and AI-driven summarization, a unified application offering these functionalities remains absent. Most existing solutions are either fragmented or limited by platform-specific constraints. This project aims to create a versatile platform that addresses these gaps, providing users with an all-in-one solution for document management and accessibility.

III. PROPOSED METHOD

A. Technology Stack

A strong technological stack was used in the development of the program to guarantee cross-platform compatibility and feature efficiency. Key components include:

- Framework: Flutter for consistent UI and performance across Android, iOS, and the web.
- State Management: Provider for optimized state handling and reactive updates.
- AI Tools: Integration of pre-trained AI models for OCR, summarization, and transcription.
- Cloud Infrastructure: Firebase for authentication and cloud storage.

B. Core Features

1) Document Conversion

- Image-to-PDF converter
- Word-to-PDF converter (vice versa)
- Excel-to-PDF converter (vice versa)
- PPT-to-PDF converter (vice versa)
- Merge PDF files

2) Optical Character Recognition (OCR)
- Extracts and processes text from images with high accuracy.

3) Speech-to-Text
- Converts spoken content into editable text.

4) AI-Driven Summarization
- Produces concise summaries of lengthy text inputs.

5) AI Tools Access & Discovery
- Enables users to effectively find and use AI-driven features by giving them access to many AI tools for different activities through a single search mechanism.

C. Design Principles

1) Responsive Design: Guarantees usability on a range of device sizes.
2) User-Friendly Interface: Designed to be easy to use and navigate.
3) Modular Architecture: Facilitates scalability and easy maintenance.

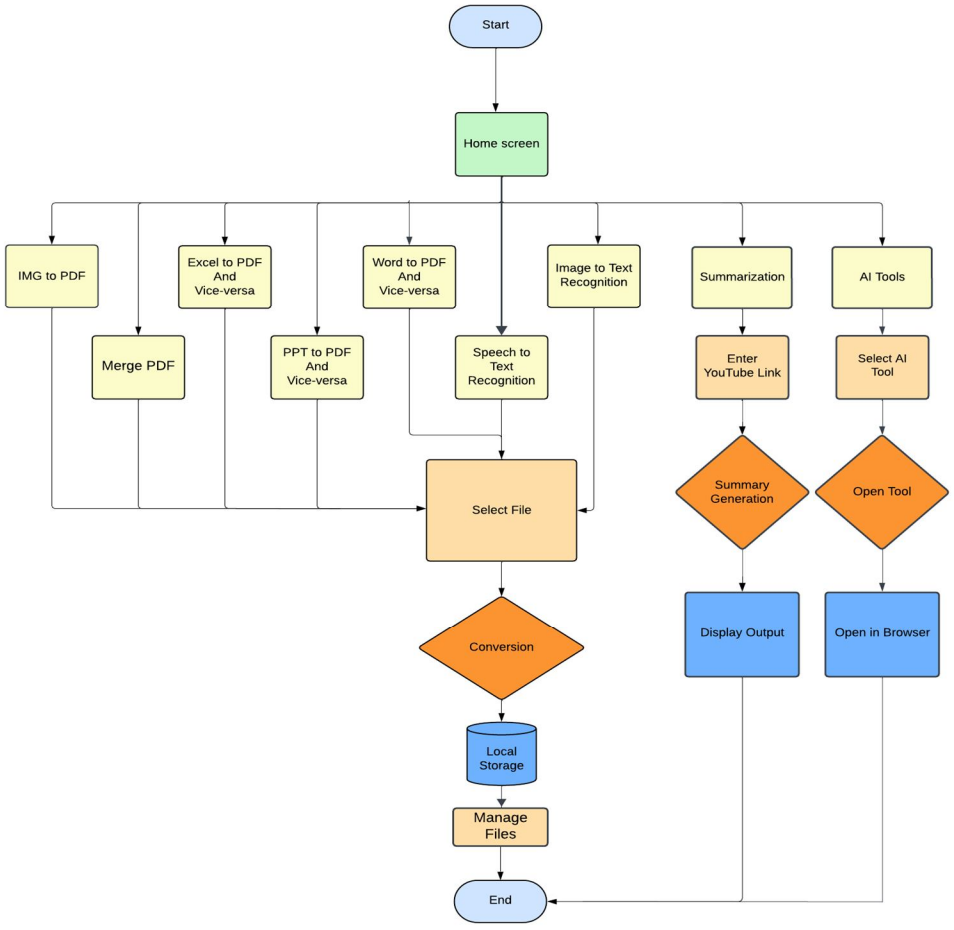


Fig 1: Block Diagram

D. Key Features and Module

1) Multi-format Document Conversion

The application supports a wide array of document format conversions, including:

- PDF to Word and Word to PDF: Facilitates easy editing and sharing of documents by enabling seamless format changes.
- PPT to PDF and PDF to PPT: Ensures users can transition between presentation and document formats without losing content fidelity.
- XLS to PDF and PDF to XLS: Supports spreadsheet management and sharing across platforms.

These features address the common challenge of format incompatibility, streamlining workflows for professionals, educators, and students.

2) Image to PDF Conversion

Users can compile multiple images into a single PDF file, ideal for archiving, sharing, and presenting visual data. Students writing project reports, professionals handling scanned documents, and everyone else trying to effectively arrange visual material will find this capability very helpful.

3) Optical Character Recognition (OCR)

Users can extract editable text from images due to the integration of OCR technology. This function is quite beneficial for:

- Conversion of printed material into digital format.
- Automating data entry from scanned documents.
- Enhancing accessibility by converting images into readable and searchable text.

4) Speech-to-Text Transcription

Harnessing speech recognition technology, the app enables real-time transcription of audio into editable text. This feature caters to:

- Professionals needing meeting notes.
- Students recording lectures.
- Individuals with accessibility needs.

5) AI-Powered Summarization

The app leverages AI to summarize textual content, including:

- Generating concise points from lengthy articles or documents.
- Creating video summaries from YouTube links. This feature enhances productivity by reducing the time spent on information processing while maintaining content integrity.

6) PDF Tools Suite

Additional tools include:

- PDF Merging: Combines multiple PDF files into a single document for better organization.
- PDF Splitting: Separates pages from a PDF file to create new documents. These tools are particularly useful for legal, academic, and business professionals managing extensive documentation.

7) Cross-platform Compatibility

The application was built using Flutter and can function smoothly on Android, IOS and web browsers. The responsive design ensures:

- Layouts that can adjust to various different screen sizes.
- Optimized performance on each platform.

8) User-Centric Design

The application incorporates:

- Intuitive navigation through a drawer menu.
- State management via Provider for smooth transitions and minimal lag.
- Customizable settings to tailor the app experience to individual needs.

E. Implementation

1) Development Environment

The FileSmith application was developed using Flutter, a cross-platform framework that ensures consistent performance across Android, iOS, and web platforms. The project follows a modular approach, integrating various document processing and AI tools. The Provider state management system is utilized to handle navigation and state efficiently.

2) Features Implemented

a) Image to PDF Conversion

- New images can be selected by users from the gallery or can be captured using the camera.
- Selected images are reorderable, allowing users to adjust their sequence before generating the PDF.
- The pdf package is used to convert images into a multi-page PDF, ensuring compression and optimized rendering.
- Users can name the PDF file and select a custom directory for saving using Flutter File Dialog.
- State management with Provider ensures smooth updates without excessive UI rebuilds.

b) PDF Merge

- Users can select multiple PDF files from storage.
- The pdf_manipulator package merges PDFs efficiently, ensuring page order integrity.
- The merged file is saved in a user-specified directory, with an option for custom naming.
- Temporary files are automatically cleared to optimize storage usage.
- Asynchronous processing ensures the app remains responsive during merging operations.

c) AI Tools Integration

- A dedicated AI Tools section provides access to various AI-powered applications.
- Users can search for tools by name or keywords, with debouncing implemented to optimize search performance.
- Each tool includes a brief description, and users can redirect to the tool's website via a built-in browser link.
- Firestore integration allows dynamic updates to the AI Tools list without app updates.

d) Speech to Text

- Developed a feature that converts spoken words into editable text using the speech_to_text package.
- Integrated an option to modify the transcribed text before saving.
- Provided the ability to save the final text as a PDF file for easy storage and sharing.

e) Extract Text from Images

- Implemented text extraction using Google ML Kit, allowing users to retrieve text from images.
- Enabled users to capture images from the camera or select from the gallery for text recognition.
- Added functionalities like text editing, text-to-speech, copying to clipboard, and PDF export.

f) PDF to DOC Conversion

- Designed a server-based approach for converting PDF to DOCX format efficiently.
- Allowed users to select a PDF, process it online, and download the converted document.
- Integrated a file-saving option for users to store the converted document at their preferred location.

3) User Interface and Experience

a) The home screen consists of feature blocks for Image to PDF, PDF Merge, and AI Tools.
b) A bottom navigation bar is implemented for intuitive navigation.
c) A menu drawer is available to access additional settings and help.

4) File Handling and Storage

a) Permissions Management: The permission_handler package is used to request file storage permissions.
b) Storage Handling: Converted and merged files are stored in a designated directory, accessible via the app.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

IV. RESULT

The performance of FileSmith was assessed based on key metrics such as conversion accuracy, processing speed, storage efficiency, and user interaction time. The evaluation was conducted on an Android device with 4GB RAM and a mid-range processor to reflect real-world performance.

A. *Speech-to-Text Conversion*

1) *Accuracy:* The system achieved an average accuracy rate of 92% in recognizing spoken words, ensuring reliable transcription.
2) *Processing Speed:* It processed speech-to-text conversion at an average of 1.2 seconds per sentence.
3) *Error Rate: An 8% error* rate was observed, primarily due to background noise and unclear pronunciation.

B. *Image-to-Text Extraction*

1) *Recognition Accuracy:* Successfully extracted 96% of printed text using Google ML Kit.
2) *Processing Time:* Each image was processed within an average of 1.5 seconds.
3) *Storage Optimization:* The extracted text was saved as a PDF, reducing storage requirements by approximately 60% compared to storing images.

C. *PDF to DOC Conversion*

1) *Conversion Accuracy:* Text-based PDFs retained 97% accuracy, while PDFs containing images and complex formatting achieved 90% accuracy.
2) *Processing Speed:* A five-page PDF was converted to DOCX format in an average of 6.2 seconds.
3) *File Size Reduction:* The converted DOCX files were 30% smaller than their original PDF counterparts, improving storage efficiency.

D. *PDF Merging*

1) *Processing Speed:* Merging a 10-page PDF took an average of 3.8 seconds.
2) *File Optimization:* The merged PDFs maintained high compression, reducing storage space by approximately 40% compared to manually combining separate files.

E. *Image to PDF Conversion*

1) *Conversion Time:* Five images were combined into a single PDF in an average of 5.6 seconds.
2) *Storage Efficiency:* Optimized compression reduced the final PDF size by 35% while maintaining readability.

F. *AI Tools Access*

1) *Search Efficiency:* Users were able to locate AI tools 30% faster using the integrated search system.
2) *Ease of Use:* The platform eliminated the need for multiple searches, providing instant access to AI-powered services like ChatGPT, DALL·E, and Google Bard.

These findings demonstrate that FileSmith delivers an effective and user-friendly solution for document conversion and management, balancing speed, accuracy, and storage optimization.

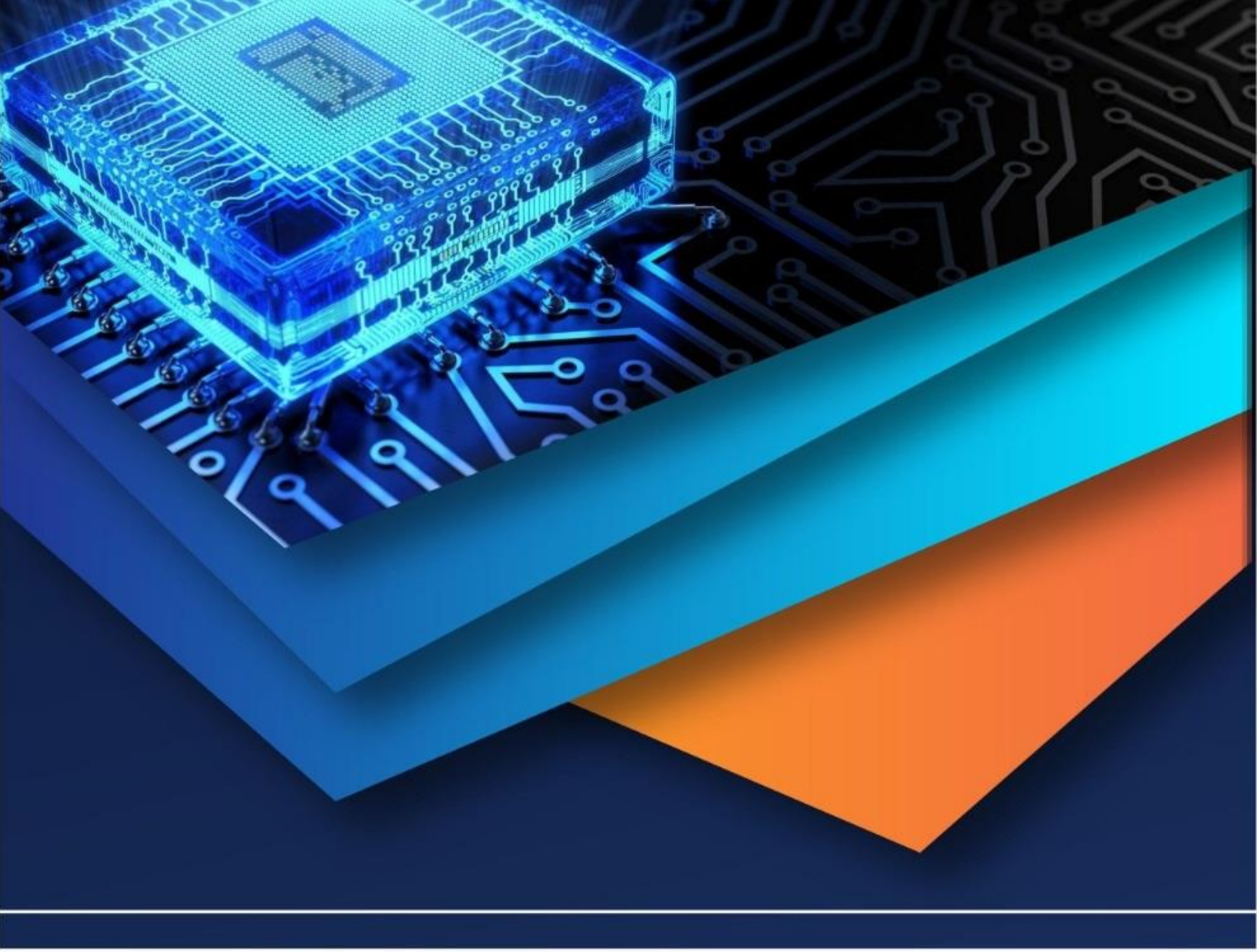
V. CONCLUSION

This research presents a flexible document conversion app developed with Flutter, designed to meet modern productivity demands. With a user-friendly interface, responsive design, and AI-powered features, it provides a seamless and efficient document management experience.

Unlike traditional tools that require multiple applications for different tasks, this app brings everything together—document conversion, OCR, and text summarization—into one unified platform. Its cross-platform support and optimized performance make it accessible to a broad audience, from students and researchers to working professionals. By leveraging AI-driven automation, the app streamlines workflows and boosts efficiency. Future enhancements, such as integrating large language models, aim to further expand its capabilities and usability.

REFERENCES

[1] Bhasin K, Goel A, Gupta G, Singh SK, Bhowmick A. A secure mobile application for speech to text conversion using artificial intelligence techniques. InWINS/CVMLH 2023 (pp. 44-53).
[2] Bagal, Vipul, Kiran Gaykar, and Ms Purnima Ahirao. "Image based Text Translation using Firebase ML Kit." Grenze International Journal of Engineering & Technology (GIJET) 8.2 (2022).
[3] Abbas, Anas, et al. "A simple PDF Converter using Android with built in editing." JOURNAL OF COMPUTER SCIENCE (ISSN NO: 1549-3636) 15.09 (2022).
[4] Penny, LaToyia DeVonne. Design & implementation of a PDF to Excel conversion tool (P2X). MS thesis. Oklahoma State University, 2008.
[5] Sahoo, Rohit, et al. "Auto-Table-Extract: A System To Identify And Extract Tables From PDF To Excel." Int. J. Sci. Technol. Res 9 (2020): 217.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY