iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52307

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 11 Issue V May 2023- Available at www.ijraset.com

Superpdf: One-Stop Solution to All Your PDF Problems

Harit Joshi[1], Mayank Belwal[2], Abha Sharma[3]
[1]*Dept. of IT, Inderprastha Engineering College*
[2]*Dept. of IT, Inderprastha Engineering College*
[3]*Assistant Professor, Dept. of IT, Inderprastha Engineering College, UP, India*

Abstract*: Superpdf is a web-based application that facilitates the modification and manipulation of PDF files. SuperPDF lets users combine multiple PDFs into one, split a PDF into multiple pages, convert PDFs to other formats, and change PDF content by adding text, photos, and notes. One of the primary advantages of SuperPDF is its user-friendly interface, which makes it simple for users of all ability levels to explore and utilize the software's different capabilities. In addition to a variety of advanced features, such as the ability to password-protect PDFs and redact sensitive information, the tool provides a variety of other capabilities. In addition to its core PDF editing capabilities, SuperPDF offers a variety of additional tools and functions, including the ability to build and fill out PDF forms, extract images and text from PDFs, and add electronic signatures to documents.

Index Terms: JBIG2, lossless, bi-level image compression, OCR, image compression, MMR, model-based coding, Tesseract, Encryption

I. INTRODUCTION

SuperPDF is a potent application for working with PDF files. It is a web-based platform for editing, converting, merging, and splitting PDF documents.

It's designed to make it simple for people to alter PDF files without the need for costly software or complex technical knowledge. It boasts a straightforward, user-friendly design that enables users to complete several tasks fast and without difficulty.

Some of its important features include

1) *PDF Edit:* It permits users to edit PDF files by adding or removing text, graphics, and other features. Users can also rotate, trim, and resize pages, as well as add or remove pages from a document.
2) *PDF Conversion:* It can convert PDF files to Microsoft Word, Excel, and PowerPoint, as well as JPG and PNG picture formats. It can also convert Word and Excel documents to PDF format.
3) *PDF Splitting:* SuperPDF can divide PDF files based on page range, page count, or file size. It is also capable of splitting by bookmarks and text occurrences.
4) *PDF Encryption:* SuperPDF permits users to password-protect PDF files in order to prevent unauthorized access. It also supports the addition of digital signatures to PDF documents.
5) *PDF Merging:* It enables users to combine numerous PDF files into one. It may combine PDFs in numerous ways, including by page range, page order, and file size.

SuperPDF is accessible over the web and as a desktop application for Windows, Mac, and Linux. It is also available as an iOS and Android mobile application. Its adaptability is one of its primary advantages. Individuals, small enterprises, and huge organizations can utilize this product. It is also appropriate for users of all ability levels, from novices to experts.

Another benefit of SuperPDF is its reasonable price. It has a variety of payment choices, including a free plan that permits users to complete a restricted amount of activities. Paid plans are also available for users who require more advanced features or greater usage restrictions.

Overall, Superpdf is a comprehensive, user-friendly, and affordable platform that provides users with the tools they need to manage and edit their PDF files. Whether you are a student, professional, or small business owner, Smallpdf offers a range of features and benefits that make it an excellent choice for anyone looking for a convenient and reliable PDF management solution.

©IJRASET: All Rights are Reserved

II. LITERATURE REVIEW

This paper provides an overview of the key components of the Tesseract OCR engine, including the image pre-processing steps, feature extraction techniques, and machine learning algorithms used for classification.

The authors also discuss the various language models and dictionaries that can be used with Tesseract to improve its accuracy on different languages and scripts.

The paper describes several key features of Tesseract, including its ability to recognize multiple fonts and text sizes, its support for various image file formats, and its adaptability to different languages and scripts.

The authors also discuss some of the challenges of OCR, such as noise and distortion in the image, and describe the techniques used by Tesseract to address these challenges.

It presents an evaluation of the performance of Tesseract on a variety of datasets and compares its accuracy to other OCR engines.

It is sometimes referred to as Group 4 or T.6 faxing.

Fax G4 is the most popular fax standard, and most fax machines and software support it. (Feng, Fuchs, & Bouman, 2003) Modified Huffman coding is a lossless compression algorithm used by Fax G4 to reduce the amount of data delivered while keeping the integrity of the original document.

It is faster than other fax transmission standards, as it can transmit up to 24 pages per minute.

Fax G4 is used in a variety of situations, including businesses, healthcare facilities, and government agencies.

Text and line art documents are often easier to compress with the Modified Huffman coding scheme, making them ideal for transmission.

Efficiency is one of the key features of Fax G4.

It can send a huge amount of data in a relatively short period of time, making it an efficient and cost-effective way for document transmission.

Fax G4 is also extremely simple to use, given the majority of fax machines and software support it.

Fax G4 is a fax transmission standard used to send faxes over telephone lines. It is sometimes referred to as Group 4 or T.6 faxing.

Fax G4 is the most popular fax standard, and most fax machines and software support it.

JBIG2 will allow both quality-progressive coding through refinement stages, with the progression going from lower to higher quality, and content-progressive coding, successively adding different types of image data. Lossy compression is usually employed for scanned documents that are largely text.

JBIG2 will allow both quality-progressive coding through refinement phases, progressing from lower to higher quality, and content progressive coding, adding different forms of image data.

Since a digital signature is just a sequence of zeroes and ones, it is desirable for it to have the following properties: the signature must be a bit pattern that depends on the message being signed (thus, for the same originator, the digital signature is different for different documents); the signature must use some information that is unique to the sender to prevent both forgery and denial; it must be relatively easy to produce; it must be relatively easy to recognize and verify the authenticity of digital signature; it must be computationally infeasible to forge a digital signature either by constructing a new message for an existing digital signature or constructing a fraudulent digital signature for a given message; and it must be practical to ret-copies of the digital signatures in storage for arbitrating possible disputes later.

However, message authentication techniques cannot be directly used as digital signatures due to inadequacies of authentication techniques. In practice, instead of using the whole message, a hash function is applied to the message to obtain the message digest.

Thus, as the number of user pairs Digital signatures increases, it becomes extremely difficult to generate, distribute, and keep track of the secret keys.

We did a lot of online research to develop a list of current PDF editors in order to obtain information for this poll. The features and functionality of each editor were then assessed using a mix of user evaluations, web reviews, and our own testing. We also took into account each editor's cost, usability, and general performance.

Additionally, we researched the numerous PDF editors that are currently on the market. There are now more PDF editors available than ever before, each with its own special features and skills. In this essay, we set out to give a thorough analysis of the many PDF editors on the market and assess their advantages and disadvantages.

III. METHODOLOGY

A. Requirements

Table 1: Requirements

HARDWARE	SOFTWARE
2GB Ram	Browser supported: Chrome (> 56.0.3239.132) Edge (> 36.16299.15.0) Safari (> 8.0.2)
> Pentium processor, intel i3 processor	JavaScript must be enabled
Internal GPU including screen size < 32 inches	Operating System: > Windows 7 > Mac OS Mojave >Ubuntu 14.04

B. Frameworks Used

ReactJS
NodeJS
MongoDB
TypeScript
Mongoose ORM
ExpressJS
FireBase
PDFtron API
Material UI

Superpdf is a cloud-based online platform that provides a suite of tools for editing, converting, compressing, and merging PDF files. The following methodology outlines the steps taken by Superpdf to deliver its services to its users.

1) *User Upload:* Users can upload PDF files to the Superpdf platform from their computer, Google Drive, Dropbox account, using a secure and encrypted connection.
2) *File Processing:* Superpdf performs a processing stage to ensure the file is ready for editing, converting, compressing, or merging.
3) *File Conversion:* Superpdf uses proprietary algorithms to convert PDF files into a variety of formats, including Microsoft Office, image formats.
4) *File Editing:* It provides editing tools, images, shapes, and lines to PDF files.
5) *File Compression:* It is a compression tool that reduces the size of a PDF file without affecting its quality.
6) *File Merging:* Superpdf provides a merging tool to merge multiple PDF files into one.

C. Implementation

Superpdf offers a variety of PDF editing tools, including the ability to split, merge, and rotate PDFs. It also provides PDF conversion services to enable users to convert their PDF files into other file formats. Finally, users can easily sign and send documents for electronic signature using Superpdf's PDF signature feature.

A PDF security feature from Superpdf enable users to password-protect their PDFs and manage access to private documents.
A PDF optimization feature from Superpdf might enable users to compress their PDF files for quicker downloads and sharing.
A PDF annotation feature from Superpdf might enable users to annotate their PDF files with notes, comments, and highlights.
A PDF OCR (Optical Character Recognition) feature from Superpdf might enable users to take the text out of scanned documents and images.

Different modules involved in the project are

1) *Editing PDFs:* With the help of this module, users will be able to add or remove text, images, and other elements from their PDF documents.
2) *PDF Merge:* Users of the PDF Merging module could combine multiple PDF files into a single document. It might offer options for changing the layout of the pages and the order in which they appear.
3) *PDF Splitting:* With the help of this module, users could divide a single PDF file into multiple documents based on the number of pages or the file size.
4) *PDF Signing:* Using this module, users could digitally sign PDF files and send them to others to sign. It might have options like setting the signing order and adding a visible signature image.
5) *PDF Security:* Users would be able to password-protect their PDF files and manage access. It might have options for controlling printing, copying, and editing permissions as well as to add digital signatures.
6) *PDF Compression:* With the help of this module, users could shrink the size of their PDF files by compressing the images and other document components. This help to speed up downloads and use less storage.

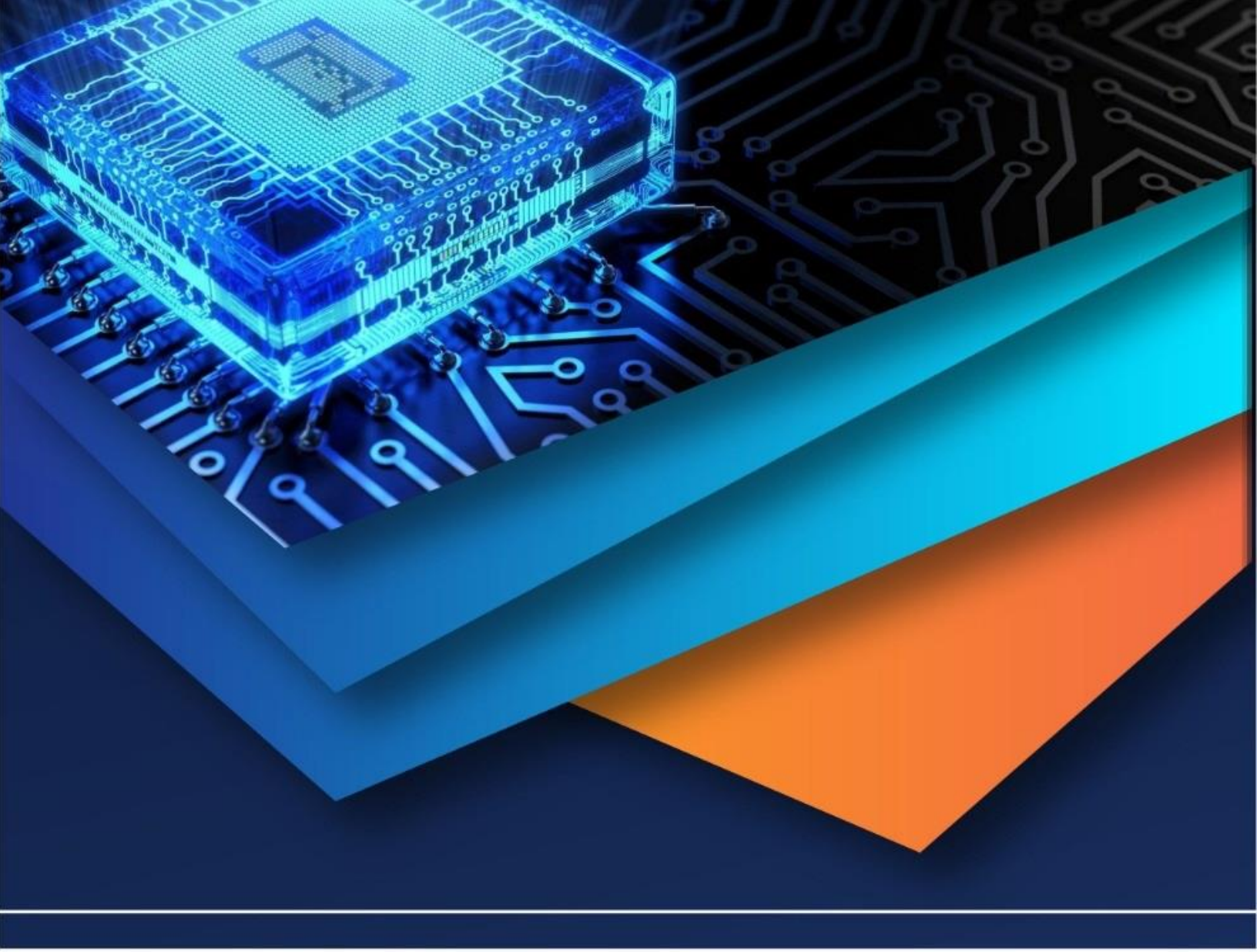
IV. DISCUSSION AND CONCLUSION

In conclusion, Superpdf is a cloud-based online platform that provides a suite of tools for editing, converting, compressing, and merging PDF files. The platform's user-friendly interface and secure connection make it a popular choice for millions of users worldwide. With its proprietary algorithms and a focus on security and efficiency, Superpdf is a reliable and convenient solution for managing PDF files. Whether it's converting a file into a different format, compressing a file to reduce its size, or merging multiple files into one, This makes Superpdf a valuable resource for anyone who frequently works with PDF files.

REFERENCES

[1] https://ieeexplore.ieee.org/abstract/document/6205503 - I. F. A. Shaikhli, A. M. Zeki, R. H. Makarim and A. -S. K. Pathan, "Protection of Integrity and Ownership of PDF Documents Using Invisible Signature," 2012 UKSim 14th International Conference on Computer Modelling and Simulation, Cambridge, UK, 2012, pp. 533-537, doi: 10.1109/UKSim.2012.81.
[2] Ghosh, K., Chakraborty, A., Parui, S.K., Majumder, P.: Improving information retrieval performance on OCR text in the absence of clear text ground truth. Inf. Process. Manag. 52(5), 873–884 (2016) https://www.sciencedirect.com/science/article/abs/pii/S030645731630036X?via%3Dihub
[3] Adobe Developer Connection: PDF Reference and Adobe Extensions to the PDF Specification. (n.d.). Retrieved 2 15, 2023, from Adobe Systems: https://www.adobe.com/devnet/pdf/pdf_reference.html
[4] B, J., & Rao, A. (2013). An approach for Image Compression Using Adaptive Huffman Coding. International journal of engineering research and technology, 2(12). Retrieved 2 15, 2023, from https://ijert.org/research/an-approach-for-image-compression-using-adaptive-huffman-coding-ijertv2is121242.pdf
[5] Borovikov, E. (2014). A survey of modern optical character recognition techniques. arXiv: Computer Vision and Pattern Recognition. Retrieved 2 15, 2023, from https://arxiv.org/pdf/1412.4183
[6] Burr, W. (2003). Selecting the Advanced Encryption Standard. IEEE Security & Privacy, 1(2), 43-52. Retrieved 2 15, 2023, from https://dl.acm.org/citation.cfm?id=859069
[7] Constantin, L. (n.d.). Microsoft continues RC4 encryption phase-out plan with .NET security updates. Retrieved 2 15, 2023, from http://www.computerworld.com/article/2489395/encryption/microsoft-continues-rc4-encryption-phase-out-plan-with--net-security-updates.html
[8] Feng, G., Fuchs, M. G., & Bouman, C. A. (2003). Image rendering for digital fax. electronic imaging, 5008, 504-512. Retrieved 2 15, 2023, from https://core.ac.uk/display/20960110
[9] FIPS PUB 186-3: Digital Signature Standard (DSS), June 2009. (n.d.). Retrieved 2 15, 2023, from http://csrc.nist.gov/publications/fips/fips186-3/fips_186-3.pdf
[10] Ginting, R. U., & Dillak, R. Y. (2013). Digital color image encryption using RC4 stream cipher and chaotic logistic map. Retrieved 2 15, 2023, from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6676220
[11] Howard, P., Kossentini, F., Martins, B., Forchhammer, S., & Rucklidge, W. (1998). The emerging JBIG2 standard. IEEE Transactions on Circuits and Systems for Video Technology, 8(7), 838-848. Retrieved 2 15, 2023, from https://ieeexplore.ieee.org/document/735380
[12] Kamali, S. H., Shakerian, R., Hedayati, M., & Rahmani, M. (2010). A new modified version of Advanced Encryption Standard-based algorithm for image encryption. Retrieved 2 15, 2023, from http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.ieee-000005559902
[13] Mazzeo, M. (n.d.). Digital Signatures and European Laws. Retrieved 2 15, 2023, from Symantec: http://www.symantec.com/connect/articles/digital-signatures-and-european-laws

[14] Patel, C. I., Patel, A., & Patel, D. (2012). Optical Character Recognition by Open source OCR Tool Tesseract: A Case Study. International Journal of Computer Applications, 55(10), 50-56. Retrieved 2 15, 2023, from http://assistivetechnology.vcu.edu/wp-content/uploads/sites/1864/2013/09/pxc3882784.pdf
[15] Rakshit, S., Basu, S., & Ikeda, H. (2010). Recognition of Handwritten Textual Annotations using Tesseract Open Source OCR Engine for information Just In Time (iJIT). arXiv: Computer Vision and Pattern Recognition. Retrieved 2 15, 2023, from https://arxiv.org/pdf/1003.5893
[16] Smith, C. M. (1993). Efficient software implementation of the JBIG compression standard. Retrieved 2 15, 2023, from https://scholarworks.rit.edu/theses/174
[17] Smith, R. (2007). An Overview of the Tesseract OCR Engine. Retrieved 2 15, 2023, from https://ai.google/research/pubs/pub33418
[18] Umamageswari, A., & Suresh, G. R. (2014). A New Cryptographic Digital Signature for Secure Medical Image Communication in Telemedicine. International Journal of Computer Applications, 86(11), 4-9. Retrieved 2 15, 2023, from https://ijcaonline.org/archives/volume86/number11/15027-3343
[19] Valliappan, M., Evans, B. L., Tompkins, D. A., & Kossentini, F. (1999). Lossy compression of stochastic halftones with JBIG2. Retrieved 2 15, 2023, from http://ieeexplore.ieee.org/document/821600



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)