iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** II **Month of publication:** February 2023

DOI: https://doi.org/10.22214/ijraset.2023.49108

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 II Feb 2023- Available at www.ijraset.com

Revolutionizing Data Entry: An In-Depth Study of Optical Character Recognition Technology and Its Future Potential

Aaryan Raj [1], Sakshi Sharma [2], Janvee Singh [3], Aastha Singh[4]
[1, 2, 3]*Department of Engineering-Computer Science, Chandigarh University, Mohali, Punjab-140413*
[4] *School of Computer Engineering & Technology, MIT World Peace University, Pune, Maharashtra, India-411038*

Abstract: ***This research paper delves into the advancements and potential of Optical Character Recognition (OCR) technology for revolutionizing data entry processes. The study provides a comprehensive analysis of the history, current state, and future of OCR technology. The paper begins with an overview of the evolution of OCR technology and its various applications in document scanning, indexing, and management.***
The research then examines the benefits of OCR technology, such as improved accuracy, faster processing times, and cost savings compared to traditional data entry methods. It also addresses the challenges faced by OCR technology, including difficulties in recognizing handwritten text and the need for continuous improvement and innovation. The paper concludes with a discussion of the future potential of OCR technology, including the integration of artificial intelligence and machine learning for improved accuracy and efficiency, and the potential for increased automation in various industries. The study highlights the importance of OCR technology for revolutionizing data entry processes and its potential for further development and growth in the future.
Keywords: OCR, Artificial Intelligence, Document Scanning, Machine Learning, Image Recognition

I. INTRODUCTION

In today's digital world, the vast amounts of data being generated on a daily basis require efficient and accurate methods of data entry. One technology that has been instrumental in revolutionizing data entry is Optical Character Recognition (OCR). OCR is the process of converting images of typed or handwritten text into editable and searchable digital text. With the advent of deep learning algorithms, the accuracy and speed of OCR technology have improved dramatically, making it a crucial tool in many industries, such as healthcare, finance, and retail.

However, there are still several challenges that need to be addressed to further improve the accuracy and reliability of OCR technology, particularly when it comes to recognizing text in low-resolution images or handwritten text. This research paper aims to provide an in-depth study of OCR technology, including its current state-of-the-art and its future potential, as well as a comprehensive examination of the challenges and solutions for OCR in different contexts and applications.

By exploring the impact of deep learning algorithms, the potential for error correction, and the ethical and privacy implications of OCR technology, this research paper aims to shed light on the future potential of this revolutionary technology in data entry. Additionally, the paper will also explore the potential for integrating OCR technology with other technologies, such as natural language processing, to improve data entry and enable new applications. Furthermore, the paper will conduct a comprehensive survey of the current and potential commercial applications of OCR technology, and identify the industries with the highest potential for OCR technology adoption.

This will provide valuable insights into the current and future uses of OCR technology, and help to identify new and innovative opportunities for its implementation. Moreover, the paper will also provide insights into the future trends in OCR technology, such as the development of more accurate and sophisticated deep learning algorithms, and the integration of OCR technology with other technologies such as blockchain and the Internet of Things (IoT). This will provide a glimpse into the future potential of OCR technology, and help to identify new and innovative opportunities for its implementation.

Additionally, the paper will also delve into the ethical and privacy implications of OCR technology, such as the security of personal data and the potential for misuse. This will help to raise awareness of the potential risks associated with OCR technology and to ensure that the development and implementation of OCR technology is done in a responsible and ethical manner.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue II Feb 2023- Available at www.ijraset.com

II. HISTORICAL OVERVIEW OF OCR TECHNOLGY

Optical Character Recognition (OCR) technology has revolutionized the way data is entered into computers and digital systems. The technology allows for the automated conversion of scanned images and scanned documents into machine-readable text, significantly reducing the time and effort required for manual data entry.

According to H. F. Schantz (Author: The History of OCR, Optical Character Recognition), OCR's roots can be traced back very far. In fact, the first patent related to OCR was issued in 1809. Around 1870 Boston's Charles R. Carey patented an image transmission system using photovoltaic mosaics. This was an early example of a "retinal scanner". Charles R. Carey invented the retinal scanner, an image transmission system that uses mosaics of photocells . Twenty years later, German inventor Paul Nipkow developed another image scanner that gave a decisive impetus to modern televisions, readers, and character recognition. A remarkable leap forward took place in 1890 with the advent of modern television sets and viewing machines featuring Nipkow's innovation, a series of scanners. In OCR, an early age was regarded as helping the blind.

Emanuel Goldberg was indeed a key figure in the early development of OCR technology. At the time of first world war he invented a machine that could read characters and convert them into telegraph code. As a physicist and inventor, he recognized the potential of using machines to automate the process of converting written text into machine-readable code. His invention of a machine that could read characters and convert them into telegraph code was a significant step forward in the development of OCR technology, and laid the foundation for further advancements in the field.Goldberg's work was a precursor to the development of OCR technology, which would later be used for a variety of purposes, including data entry, document digitization, and text recognition. The impact of his work has been far-reaching, and OCR technology continues to play an important role in the digitization of industries and the automation of data entry processes. In 1951, American inventor David Hammond Shepard, a cryptanalyst for the Armed Forces Security Administration (AFSA), predecessor of the National Security Agency (NSA), built Gismo in his spare time. "Gismo" is a machine that converts printed messages into machine language for computer processing and is the first Optical Character Recognition (OCR) system. The machine is known primarily from United States Patented S. 2,663,758O was filed by Shepard on March 1, 1951 and patented on December 22, 1953. This patent is entitled Reading Apparatus. The present invention relates to a method and apparatus for analyzing information and the like. Briefly, the present invention relates to a so-called reader designed to recognize printed characters, punched holes, etc. and to recognize the identity of a particular character or other element passing in front of a recognition means reproduced in various forms for coding.

OCR technology was first introduced in the 1960s, but it wasn't until the advent of computer technology in the 1980s and 1990s that it became widely used. Initially, OCR was used primarily in the banking and financial industry, where large amounts of data needed to be processed quickly and accurately. Over the years, OCR technology has continued to evolve and improve, with more advanced algorithms and machine learning techniques being developed to enhance its accuracy and reliability. Today, OCR is used in a wide range of industries, including healthcare, retail, and government, to streamline data entry processes and improve efficiency. In recent years, OCR technology has also been integrated into mobile devices and smartphones, allowing users to scan text from documents and images and convert them into machine-readable text on-the-go. This has made OCR technology even more accessible and has further increased its widespread use. Additionally, with the growth of artificial intelligence and machine learning, it is likely that OCR technology will continue to improve, becoming even more accurate and versatile. This could open up new possibilities for OCR in industries such as law, education, and more, further revolutionizing the way data is entered into digital systems.

III. IMPACT OF OCR IN DATA ENTRY PROCESSES

Optical Character Recognition (OCR) technology has had a significant impact on data entry processes. OCR is a process that uses image recognition algorithms to identify and extract text from scanned documents and images. The impact of OCR technology on data entry processes has been significant and far-reaching. Some of the key ways in which OCR has revolutionized data entry include:

1) *Increased Speed And Efficiency:* OCR technology enables the automated conversion of written text into machine-readable code, reducing the time and effort required for manual data entry.
2) *Improved Accuracy:* OCR technology is able to recognize characters with high accuracy, reducing the risk of errors and inaccuracies in the data entry process.
3) *Reduced Costs:* OCR technology can significantly reduce the cost of data entry processes, as it eliminates the need for manual labour and reduces the risk of errors and inaccuracies.

©IJRASET: All Rights are Reserved

4) *Increased Accessibility:* With the introduction of OCR software for personal computers and mobile devices, the technology has become more accessible to a wider range of users, including small businesses and individuals.
5) *Improved Data Organization And Management:* By automating the data entry process, OCR technology can help organizations to better organize and manage their data, making it easier to access, analyse, and utilize.
6) *Increased Opportunities For Data Analysis:* By converting written text into machine-readable code, OCR technology enables organizations to analyse and make sense of vast amounts of data in new and innovative ways.
7) *Increased Accessibility To Historical Documents:* OCR technology has made it possible to digitize historical documents and make them more widely accessible to researchers, historians, and the general public.
8) *Improved Records Management:* In many industries, including healthcare, finance, and government, OCR technology has been used to automate the process of maintaining and organizing records. This has improved the accuracy and reliability of records and has made it easier to access and analyse the data they contain.
9) *Enhanced Automation Capabilities:* OCR technology has enabled the automation of a wide range of tasks, from invoicing and receipt processing to the scanning and indexing of legal documents.

Overall, OCR technology has revolutionized the data entry process, enabling businesses to save time, improve accuracy, and increase efficiency. By digitizing paper documents, OCR technology has made it easier for businesses to access and manage their data, improving the overall management of data.

IV. LIMITATIONS AND CHALLENGES OF CURRENT OCR TECHNOLOGIES

OCR (Optical Character Recognition) technology has revolutionized the data entry process by automating the conversion of scanned documents into editable text. However, despite its many advantages, OCR technology still faces several limitations and challenges that impact its accuracy and versatility.

Table 1: Showing some limitations with OCR Technology

Limitation	Description
Handwriting recognition	OCR technology struggles with recognizing handwriting, especially if the handwriting is not legible or if it uses non-standard writing styles.
Complex document structure	OCR technology may have trouble recognizing the structure of complex documents, such as multi-column text, tables, and graphs. This can result in errors in the extracted data or loss of important information.
Non-uniform text	OCR technology may have difficulties in recognizing text that is written in different styles, font sizes, or orientations within the same document.
Image quality	The quality of the image being processed can greatly impact the accuracy of OCR technology. Factors such as image resolution, lighting, and shadows can all affect the recognition of text.
Multilingual recognition	OCR technology may struggle with recognizing text in multiple languages, especially if the text is mixed or if it uses non-standard writing systems.
Integration with other technologies	Integrating OCR technology with other technologies, such as machine learning and data management systems, can also present challenges. This may include compatibility issues, data privacy concerns, and the need for specialized knowledge and expertise.

OCR (Optical Character Recognition) technology has greatly improved the speed and efficiency of data entry processes, but it still faces several challenges that impact its accuracy and versatility. One of the main challenges of current OCR technology is text recognition accuracy. Despite advancements in OCR technology, errors can still occur during the recognition process, which can result in inaccuracies in the extracted data and require manual correction. This can be time-consuming and expensive, especially for large volumes of data.

Another challenge is the recognition of complex document structures, such as multi-column text, tables, and graphs. OCR technology may struggle to accurately recognize and preserve the structure of these documents, which can result in errors in the extracted data or loss of important information.

Additionally, OCR technology may have difficulties in recognizing text that is written in different styles, font sizes, or orientations within the same document, which can also result in errors and inaccuracies in the extracted data.

The quality of the image being processed can also greatly impact the accuracy of OCR technology. Factors such as image resolution, lighting, and shadows can all affect the recognition of text, making it difficult for OCR technology to accurately extract the data. Furthermore, OCR technology may struggle with recognizing text in multiple languages, especially if the text is mixed or if it uses non-standard writing systems.

Despite these challenges, OCR technology continues to be a valuable tool for automating the data entry process and improving the efficiency and accuracy of data management systems.

V. CURRENT APPLICATIONS OF OCR TECHNOLOGY

OCR (Optical Character Recognition) technology works by analyzing an image of text and converting it into machine-readable text. The process typically involves several steps, including pre-processing, character recognition, and post-processing.

In the pre-processing step, the image of text is prepared for analysis by removing noise, correcting the perspective, and enhancing the image quality. This helps to ensure that the text is clear and easy to read, making it easier for the OCR software to accurately recognize the characters. In the character recognition step, the OCR software uses machine learning algorithms to analyze the image and identify the characters within it. The software uses patterns and features within the image to determine the shape and position of each character, and then converts the image into machine-readable text.

In the post-processing step, the OCR software performs various checks and corrections to improve the accuracy of the output. For example, it may perform spell-checking or correction of misrecognized characters, or it may convert the text into a specific format, such as a PDF or Word document. Overall, the working model of OCR technology involves the use of machine learning algorithms to analyze images of text and convert them into machine-readable text. This process offers the potential for significant improvements in speed and accuracy over manual data entry, and has many potential applications in industries such as finance, healthcare, and publishing. However, it is important to carefully consider the ethical and data privacy implications of OCR technology and ensure that it is used in a responsible and equitable manner.

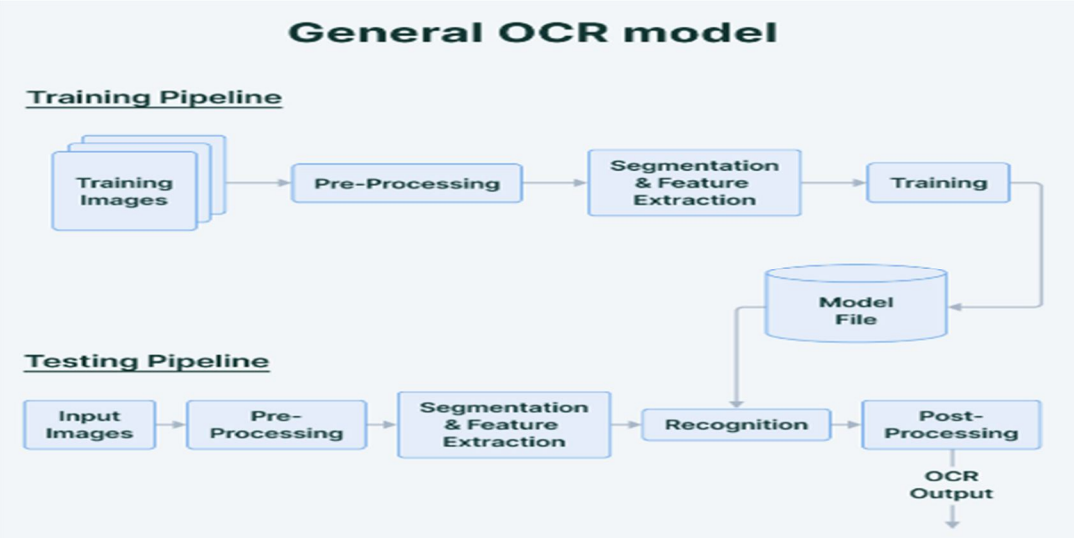


Figure 1: Working Model of OCR Technology

Optical Character Recognition (OCR) technology has become an integral part of many industries, enabling the digitization of information and streamlining data entry processes. Here are some of the unique applications of OCR technology:

1) *Document Scanning And Archiving:* OCR technology is widely used to digitize physical documents, including invoices, contracts, and government records. The technology can accurately recognize and extract text from images and convert it into editable and searchable digital documents, making it easier to store, access, and manage information.
2) *Automated Data Entry:* OCR technology is widely used in industries such as finance, healthcare, and retail to automate the process of data entry. The technology can quickly and accurately extract data from forms, receipts, and other types of documents, reducing the need for manual data entry and minimizing errors.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

3) *Business Process Automation:* OCR technology is used to automate various business processes, including invoice processing, accounts payable, and customer service. The technology can automatically recognize and extract relevant information from documents, streamlining workflows and improving efficiency

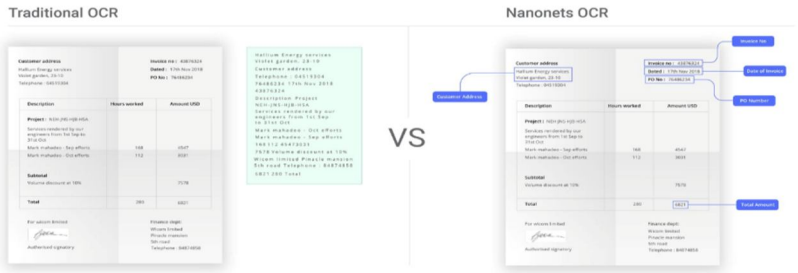


Figure 2: Showing transformation of OCR Technology

4) *Handwriting Recognition:* OCR technology is also used to recognize handwritten text, enabling the digitization of handwritten notes, journal entries, and other written records. This technology has applications in fields such as education, historical preservation, and genealogy.
5) *License Plate Recognition:* OCR technology is used in the development of automatic license plate recognition (ALPR) systems, which are used for vehicle identification and tracking, toll collection, and border security. The technology can quickly and accurately recognize license plate numbers, enabling real-time tracking and analysis of vehicle data.
6) *Book Scanning And Digitization:* OCR technology is used to digitize books, newspapers, and other types of print media, making it easier to access and preserve important information. The technology can accurately recognize and extract text from images of printed pages, allowing for the creation of searchable and editable digital versions of books and other written materials.
7) *Language Translation:* OCR technology can be used to recognize text in one language and translate it into another language, making it easier to communicate and collaborate across language barriers. This technology has applications in global business, education, and travel.



Figure 3: Showing the use of language recognition using OCR

8) *Handwritten Signature Recognition:* OCR technology is used to recognize handwritten signatures, enabling secure and convenient digital signatures in various industries, such as finance, real estate, and government. This technology eliminates the need for manual signature verification, reducing the risk of fraud and improving efficiency.
9) *Image-Based Data Extraction:* OCR technology is used to extract information from images, including barcodes, QR codes, and other types of optical data. This technology has applications in fields such as logistics, retail, and healthcare, where it can be used to automate the tracking and analysis of goods and patient data.
10) *Historical Document Preservation:* OCR technology is used to preserve historical documents, such as maps, manuscripts, and other types of written records, by converting them into digital formats. This technology enables the preservation and accessibility of important cultural heritage, improving our understanding of the past and supporting research in various fields.

VI. ETHICS OF OCR AND DATA PRIVACY

The integration of Optical Character Recognition (OCR) technology in various industries has led to improved efficiency and reduced manual errors. However, it also raises a number of ethical and data privacy concerns that must be considered. The use of OCR technology involves the processing and storage of vast amounts of personal and sensitive information, which can be vulnerable to cyber threats and data breaches. This raises important questions about who controls and owns this information, and how it is used. Additionally, the use of machine learning algorithms in OCR systems can perpetuate existing biases and discrimination, leading to inaccurate results for certain groups of people. This highlights the need for careful consideration of algorithmic bias and the need to ensure that OCR technology is implemented in a fair and equitable manner. The economic impact of OCR technology must also be taken into account, as the automation of data entry processes through OCR technology can lead to job displacement, particularly for low-skilled workers.

The lack of transparency in the inner workings of OCR systems can make it difficult to assess their accuracy and fairness, which raises questions about accountability and trust. It is essential to address these ethical and data privacy concerns through further research and the development of solutions that ensure that OCR technology is used in a responsible and ethical manner. Moreover, the use of OCR technology can raise concerns about data accuracy and completeness, as the technology is only as accurate as the data used for training. It is important to ensure that the data used for training is high quality, diverse, and representative of the population to minimize the risk of bias and discrimination. Furthermore, it is important to consider the accessibility of OCR technology, particularly for individuals with disabilities, who may require special accommodations to access and use the technology effectively. This highlights the need for inclusive design and the consideration of accessibility requirements in the development of OCR systems. It is also important to consider the regulations and laws surrounding OCR technology and data privacy. Different countries have varying laws and regulations that must be taken into account when implementing OCR technology. For example, the European Union's General Data Protection Regulation (GDPR) sets strict rules for the handling and processing of personal data, which must be followed by organizations operating in the EU.

In the United States, the Health Insurance Portability and Accountability Act (HIPAA) sets standards for the protection of sensitive medical information, which must be followed by healthcare organizations. Failure to comply with these regulations can result in significant fines and reputational damage, so it is essential to ensure that OCR technology is implemented in compliance with applicable laws and regulations. Furthermore, it is important to consider the impact of OCR technology on society as a whole. The rapid development and integration of OCR technology has the potential to greatly improve the efficiency and accuracy of data entry processes, but it must be done in a responsible and ethical manner. The use of OCR technology must be balanced with the need to protect the privacy and rights of individuals, to ensure that its benefits are realized in a way that benefits society as a whole. This requires a multidisciplinary approach that considers the ethical and social implications of OCR technology and the development of solutions that address these challenges. In summary, the integration of OCR technology into various industries presents many opportunities for improvement, but also raises important ethical and data privacy concerns that must be considered. The full implications of OCR technology must be carefully evaluated and understood, to ensure that it is used in a responsible and equitable manner that benefits society as a whole.

VII. ROLE OF MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE IN FUTURE OF OCR

The role of machine learning and artificial intelligence in the future of OCR is significant and expected to grow even more in the coming years. These technologies have the potential to greatly enhance the capabilities of OCR systems and make them even more accurate and efficient. One of the key areas in which machine learning and AI can play a role is in the character recognition step of the OCR process. By using machine learning algorithms, OCR systems can become more adept at recognizing even the most complex and difficult to read characters.

This can help to improve the accuracy of the output and reduce the number of errors that occur. Another area in which machine learning and AI can play a role is in the post-processing step. By using AI algorithms, OCR systems can become more capable of recognizing patterns and features within the text and making corrections accordingly. For example, an AI-powered OCR system may be able to recognize when a character has been misrecognized and correct it automatically, improving the accuracy of the output. In addition to these areas, machine learning and AI can also play a role in making OCR systems more flexible and adaptable. By incorporating machine learning algorithms, OCR systems can become more capable of learning and adapting to different styles of text, different languages, and different types of images. This can help to make OCR technology more accessible and useful in a wider range of industries and applications.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue II Feb 2023- Available at www.ijraset.com

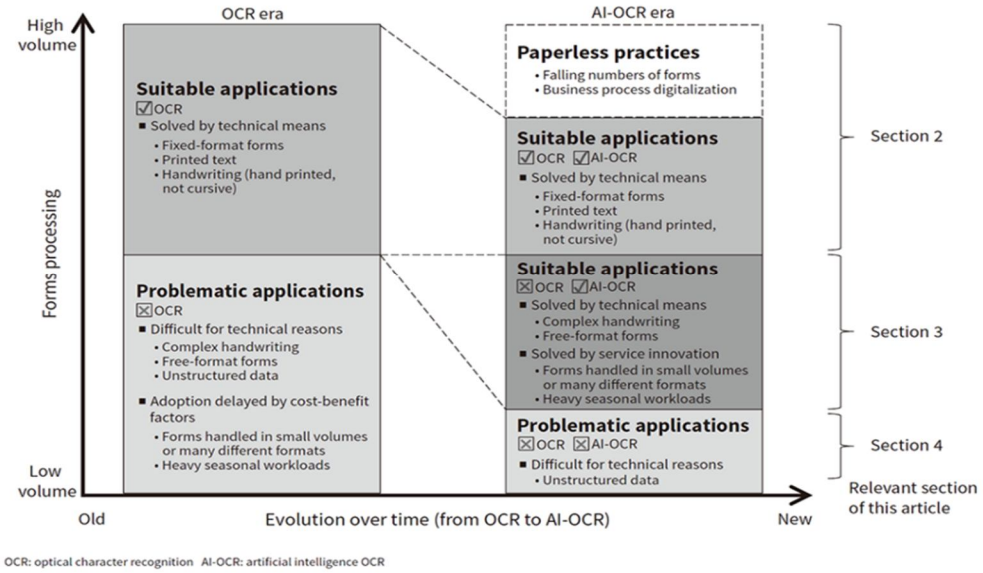


Figure 4: Showing Evolution of OCR Technology

The integration of machine learning and artificial intelligence into OCR technology has the potential to revolutionize the way that data is processed and analyzed. These cutting-edge technologies can make OCR systems faster, more accurate, and more efficient, opening up new avenues for innovation and growth.One of the most exciting possibilities for the future of OCR is the potential for real-time processing. With machine learning algorithms, OCR systems can learn from each image they process, becoming more accurate and efficient over time. This means that in the future, OCR systems may be able to process images as soon as they are captured, providing near-instantaneous results.Another area in which machine learning and AI can play a role is in improving the accuracy of OCR systems. By using advanced algorithms, OCR systems can become better at recognizing complex or unusual characters, making it possible to process a wider range of documents and images. This could be especially beneficial in industries such as legal, medical, and financial services, where precision is critical. In addition to these capabilities, machine learning and AI can also help to make OCR systems more flexible and customizable. For example, an OCR system powered by AI algorithms could be trained to recognize specific types of documents or to process images captured in different environments, such as low-light conditions. This would make OCR systems more useful and accessible to a wider range of businesses and organizations.

The integration of machine learning and artificial intelligence into OCR technology is expected to bring about several major changes and improvements in the field. These advanced technologies have the potential to enhance the accuracy, speed, and efficiency of OCR systems, making them even more valuable tools for businesses and organizations across a wide range of industries. One of the most significant benefits of incorporating machine learning and AI into OCR technology is the potential for real-time processing. With advanced algorithms, OCR systems can learn from each image they process, becoming better and more efficient over time. This could make it possible for OCR systems to process images and extract data as soon as they are captured, providing near-instantaneous results. Another key area where machine learning and AI can make a difference is in improving the accuracy of OCR systems. By incorporating advanced algorithms, OCR systems can become better at recognizing complex or unusual characters, making it possible to process a wider range of documents and images. This could be particularly useful in industries such as legal, medical, and financial services, where precision is of utmost importance.

In addition to these capabilities, machine learning and AI can also help to make OCR systems more flexible and customizable. For example, an OCR system powered by AI algorithms could be trained to recognize specific types of documents or to process images captured in different environments, such as low-light conditions. This could make OCR systems more accessible and useful to a wider range of businesses and organizations. The future of OCR technology is poised for major growth and transformation, and the role of machine learning and AI will be critical in driving this change. By harnessing these cutting-edge technologies, OCR systems will become even more powerful and versatile, enabling new applications and use cases that were previously not possible. The potential impact of these advancements on data processing and analysis is truly exciting, and we can expect to see many exciting developments in the field in the coming years.

VIII. GLOBAL MARKET FOR OCR TECHNOLOGIES

The global market for Optical Character Recognition (OCR) technologies is growing at an impressive rate, driven by the increasing demand for automation and digitalization in various industries. With its ability to extract text from images and scanned documents, OCR has become an essential tool for businesses and organizations to streamline their data entry processes, improve accuracy and efficiency, and reduce the risk of human error.According to recent market research, the global OCR market is expected to reach $3.72 billion by 2026, growing at a CAGR of 13.6% from 2020 to 2026. This growth is largely driven by the increasing adoption of OCR technology in various industries, including healthcare, finance, and government, among others.

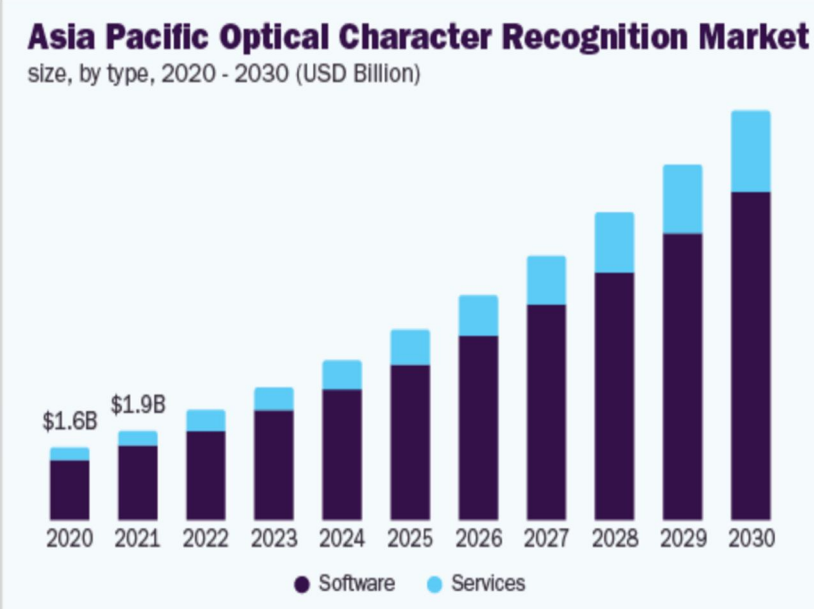


Figure 5: Showing expected OCR market till 2030

In the healthcare industry, OCR is being used to automate the process of extracting information from medical records, reducing the time and effort required for manual data entry and improving the accuracy of patient information. In the financial sector, OCR technology is being used to automate the processing of invoices, cheques, and other financial documents, reducing the risk of human error and improving the efficiency of financial transactions. In the government sector, OCR is being used to digitize and automate the process of extracting information from government records and documents, such as passport applications, voter registration forms, and census data. This not only saves time and effort, but also improves the accuracy and reliability of the data collected. Overall, the increasing adoption of OCR technology in various industries, combined with the growing demand for automation and digitalization, is expected to drive the growth of the global OCR market in the coming years. As OCR technology continues to advance and become more sophisticated, it is likely to play an increasingly important role in shaping the future of data processing and analysis. In addition to its use in various industries, the rise of mobile devices and the growing trend of paperless offices have also contributed to the growth of the OCR market. The increasing demand for mobile OCR solutions, such as smartphone-based scanning apps, is driving the development of new OCR technologies that are designed to meet the needs of mobile users.The advancement of machine learning and artificial intelligence is also playing a key role in the future of OCR technology. Machine learning algorithms are being used to improve the accuracy and speed of OCR systems, and to develop new OCR systems that can recognize a wider range of fonts and languages. With the integration of machine learning, OCR systems are becoming increasingly capable of handling complex documents, such as those with handwriting, or those with images or graphs.

Moreover, the increasing focus on data privacy and security is also driving the development of secure OCR solutions. With the growing threat of cyber-attacks and data breaches, businesses and organizations are looking for OCR solutions that can protect sensitive information and prevent unauthorized access to sensitive data.In conclusion, the global market for OCR technologies is poised for continued growth in the coming years, driven by the increasing demand for automation and digitalization, the rise of mobile devices, the integration of machine learning and artificial intelligence, and the growing focus on data privacy and security. As OCR technology continues to advance and become more sophisticated, it is likely to play an increasingly important role in shaping the future of data processing and analysis, and in revolutionizing the way we manage and analyze information.

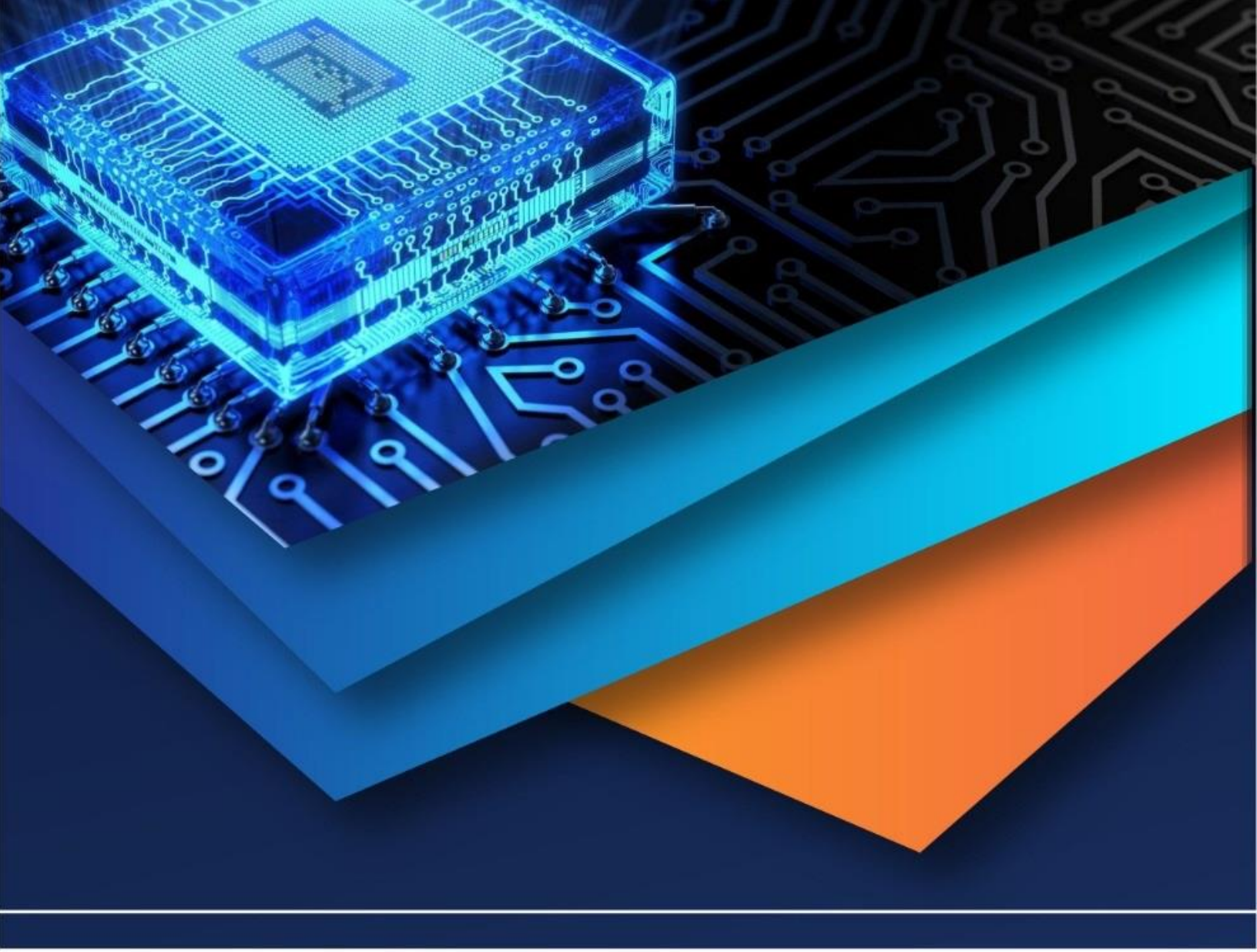
IX. CONCLUSIONS

The study of Optical Character Recognition technology has revealed its enormous potential to revolutionize data entry processes. With its ability to extract information from a wide range of images and documents, OCR has the potential to make data entry faster, more accurate, and more efficient, saving businesses and organizations time and money. The integration of machine learning and artificial intelligence into OCR technology is poised to bring about significant advancements in the field, making it possible for OCR systems to process images and extract data in real-time, improve accuracy, and become more flexible and customizable. The future of OCR technology looks bright, and its impact on data processing and analysis is likely to be substantial. However, it is important to note that there are still challenges and limitations that need to be addressed in the field of OCR. These include issues related to accuracy, document format and quality, and data privacy and security. As the technology continues to evolve, it will be crucial to address these challenges and develop solutions that ensure that OCR systems can be used in a manner that is ethical and secure. Despite these challenges, the future of OCR technology looks very promising, and its potential to revolutionize data entry processes is undeniable. As the technology continues to advance and become more sophisticated, we can expect to see many exciting developments and breakthroughs in the field, which will further enhance its impact and capabilities. As such, OCR technology is likely to play a major role in shaping the future of data processing and analysis, and its importance to businesses and organizations cannot be overstated.

REFERENCES

[1] Patel C, Patel A, Patel D. Optical character recognition by open source OCR tool tesseract: A case study. International Journal of Computer Applications. 2012 Jan 1;55(10).
[2] Ye Q, Doermann D. Text detection and recognition in imagery: A survey. IEEE transactions on pattern analysis and machine intelligence. 2015 Jul 1;37(7):1480-500.
[3] Shinde AA, Chougule DG. Text Pre-processing and Text Segmentation for OCR. International Journal of Computer Science Engineering and Technology. 2012:810-2.
[4] Jie Ding, Guotao, Fang Xu(2018) Research on Video Text Recognition Technology Based on OCR
[5] Gustav Tauschek. Reading machine. U.S. Patent 2026329, http://www.google.com/patents?vid=USPAT2026329, December 1935FLEXChip Signal Processor (MC68175/D), Motorola, 1996. [Accessed 23/11/2016]
[6] https://inkwoodresearch.com/reports/optical-character-recognition-market/
[7] https://fpt.ai/tuong-lai-cua-ocr-song-hanh-cung-ai
[8] Dinesh Acharya U, Subbareddy NV. Krishnamoorthy: Isolated Kannada Numeral Recognition Using Structural Features and K-Means Cluster. Proc. of IISN. 2007:125-9.
[9] Yetirajam M, Nayak MR, Chattopadhyay S. Recognition and classification of broken characters using feed forward neural network to enhance an OCR solution. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume. 2012 Oct 28;1.
[10] V Ohlsson(2016), Optical Character and Symbol Recognition using Tesseract
[11] Shah P, Karamchandani S, Nadkar T, Gulechha N, Koli K, Lad K. OCR-based chassis-number recognition using artificial neural networks. InVehicular Electronics and Safety (ICVES), 2009 IEEE International Conference on 2009 Nov 11 (pp. 31-34). IEEE.
[12] Mantas J. An overview of character recognition methodologies. Pattern recognition. 1986 Dec 31;19(6):425-30.
[13] Amin A. Off-line Arabic character recognition: the state of the art. Pattern recognition. 1998 Mar 1;31(5):517-30.
[14] Stallings W. Approaches to Chinese character recognition. Pattern recognition. 1976 Apr 30;8(2):87-98.
[15] Gao J, Blasch E, Pham K, Chen G, Shen D, Wang Z. Automatic vehicle license plate recognition with color component texture detection and template matching. In SPIE Defense, Security, and Sensing 2013 May 21 (pp. 87390Z-87390Z). International Society for Optics and Photonics.
[16] Mishra N, Patvardhan C. ATMA: Android Travel Mate Application. International Journal of Computer Applications. 2012 Jan 1;50 (16).

©IJRASET: All Rights are Reserved



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)