IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SURVEY ON ADVANCEMENT IN OCR-BASED TECHNOLOGY FOR SEAMLESS MULTILINGUAL COMMUNICATION

Ronit Keshave[1] ,Prof. Vina M. Lomte[2], Aditi Mooley [3], Chaitanyya Sawant [4], Shivani Kanojia [5]

[1-5]Department Of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune - 411058, India.

Abstract:– In our increasingly globalized world, effective communication is vital, yet language barriers often hinder smooth interaction between people from different linguistic backgrounds. To address this issue, technological advancements have led to the development of special sunglasses embedded with [5]Optical Character Recognition (OCR) technology, enabling real-time language recognition and translation. These special sunglasses are designed to help people communicate more easily, especially when they encounter languages they don't understand. We discuss the significance of machine learning and artificial intelligence (AI) techniques in enhancing the accuracy and efficiency of language recognition, while also addressing the complexities and limitations associated with the recognition of various fonts, styles, and handwriting. We take a close look at different methods and techniques that have been used in this field and we discussed the good things and the challenges that come with using OCR-based technology for language recognition and translation. Furthermore, we highlight the challenges associated with the integration of OCR technology into wearable sunglasses, such as ensuring real-time processing, minimizing hardware constraints, and optimizing energy consumption for prolonged usage. we provide a detailed overview of the latest and most advanced approaches being used, how they are applied in real life, and what we can expect in the future. Our goal is to help readers understand the full potential of these OCR-based sunglasses in breaking down language barriers and fostering better communication between people from different cultures.

Index Terms – OCR Technology, Effective Communication, Language Barriers, Real-time language Recognition, Translation, Accuracy, Artificial Intelligence.

I. INTRODUCTION

In the fast-paced landscape of technological innovation, the convergence of artificial intelligence and wearable devices has given rise to ground breaking solutions that redefine the way we interact with the world around us. Among these cutting-edge advancements, the integration of [5]Optical Character Recognition (OCR) technology into smart sunglasses stands out as a transformative breakthrough, promising to revolutionize the realm of language recognition and translation.

This synergy empowers users with the ability to capture and interpret textual information from their immediate surroundings, transcending language barriers in a manner that is both intuitive and unobtrusive. These smart glasses harness the power of technology to break down language barriers and enable the blind to engage in a world of diverse languages with confidence and ease.

These wearable devices represent a remarkable convergence of technology and fashion, with the potential to revolutionize the way we perceive and interact with languages. Our mission is to offer a comprehensive analysis of the present state of this technology, from the underlying technical workings to the real-world applications and the challenges that still need to be overcome. .

OCR, is a technology that has been quietly making our lives easier for years. It enables machines to recognize printed or handwritten text, a task that's second nature to humans but requires sophisticated algorithms for computers. OCR has found applications in document scanning, digitization, and text extraction from images, which means it can read text from pictures OCR technology is the cornerstone of OCR-based sunglasses. These smart shades come equipped with tiny cameras and powerful processors. They can "see" the world through their lenses, capturing text from a variety of sources, such as signs, books, menus, or any printed material. The text is then processed in real-time, recognized by the OCR technology, and translated into the language of your choice.

By providing real-time assistance in reading and comprehending text, the sunglasses can support visually impaired individuals in performing daily activities independently. Whether it is reading product labels while shopping, identifying bus schedules at public transportation stops, or understanding instructions on household items, these sunglasses can enhance autonomy and self-reliance in various aspects of daily life.

The language recognition and translation capabilities of the sunglasses can aid visually impaired individuals in learning new languages. By providing instant audio translations of foreign text, the sunglasses can assist in language comprehension, pronunciation, and vocabulary building, thereby facilitating an inclusive and accessible learning experience for language learners with visual impairments.

This technological synthesis holds the promise of transforming the way we perceive and engage with languages, forging connections that transcend linguistic boundaries. As we delve deeper into the intricacies of this groundbreaking innovation, we unravel a tapestry of possibilities that may redefine the very fabric of global communication in the years to come.

II. LITERATURE SURVEY

Ref. No.	Paper Title	Publication Details	Pre-Processing Feature	Feature Extraction and Classification	Accuracy	Post Processing	Research Gap Identified
[1]	Optical Character Recognition with Word Prediction Feature using First Order Hidden Markow For the Blind	Chew H. Wen, Lim L. Tze, 2021	HMM for Misspelling Correction	pytesseract for OCR	94.05%	pyttsx3 for Text-to-Speech	Implementation of OCR on glasses with a small camera for scanning and Bluetooth earphone for speech output
[2]	Recognition and Speech Conversion of Devnagari Script using CNN	Sameer Agrawal, Neelam Agrawal, 2023	Image Preprocessing, Data Augmentation	Convolutional and Pooling Layers, Activation Functions	90%	Thresholding, Binarization, Noise Removal, Punctuation and Grammar Correction, Text	Improvement in quality of image segmentation, image classification, recognition, and speech conversion

						Cleaning and Normalization, Contextual Analysis and Correction	

[3]	Smart Reader for Visually Impaired	Abbineni Charishma, Alla Amrutha Vaishnavi, D Rajeswara Rao, Tirumalasetti Teja Sri, 2023	Image Quality Enhancement, Noise Reduction	Feature Extraction from Characters, TTS Algorithms	92.50%	Error Correction, Language Model Integration	Image quality, te placement, and orientation
[4]	A Peculiar Reading System for Blind People using OCR Technology	M. Raja, J. Deny, Nagaraj P, Muneeswaran. V, 2023	Text Region Localization	Spatial Layout, Font Size, Text Style Features	95%	Interface Customization, Context-Sensitive Correction	Incorporati advanced NLP techniques for synthesized speech output
[5]	Multilingual Text & Handwritten Digit Recognition and Conversion of Regional languages into Universal Language Using Neural Networks	Dr. Bhushan Vidhale, Prof. Ganesh Khekare, Prof.Chetan Dhule, Dr. Pankaj Chandankhede, Prof. Abhijit Titarmare, Prof. Meenal Tayade, 2021	Various Image Preprocessing techniques	Glyph Summation	92.50%	Canvas for Text Drawing	Statistical and geometric features recognition
[6]	Text Extraction and Recognition System for Myanmar Warning Signboard Images	Kyi Pyar Zaw, Nu War, 2023	Specific Preprocessing for Myanmar Language	Chain Code, Pixel Density, Centroid Distance Features	92.77%	Character Layout Reconstruction	Robust features Myanmar script, Mul lingual character recognition
[7]	RASPBERRY PI BASED BOOK READER FOR VISUAL IMPAIRED PEOPLE	Pratiksha Kapgate, Sakshi Tidke, Rutwik Fender, Sarthak Rathore, Sumeet Ghodmare,	OCR, Text Preprocessing	Feature Extraction and Machine Learning Classification	93.20%	N/A	Advanced features a seamless integration

		Prof. Leela Bitla, 2023					
[8]	Portable Camera-based Identification System for Visually Impaired People	Ashveena A, Bala Deepika J, S. Prince Mary D. Usha Nandini, 2023	Single Shot Detector (SSD) for Image Processing	Extraction of Visual Cues	95%	Object Refinement, Feedback to Users	Use healthcare, safety, a education
[9]	Optical Character Recognition for Medical Records Digitization with Deep Learning	Muhammad Ateeque Zaryab, Chuen Rue Ng, 2023	Normalization, Noise Handling	CNN for Feature Extraction	83%	Spell Checking, Grammar Correction, Context Validation	Handling complex handwritin styles
[10]	OCR Text Region Detection Based on Chinese Character	Li Chen, Yong Qiao, Shikan Fu, Jing Cao, Lei Wang,2023	Gabor Filters for Chinese Character Features	Stroke Patterns, Structural Elements	80.30%	Non-maximum Suppression	Deep learning f Chinese character recognition

[11]	Multi-connected Tiny Object Detection for Color-Strip Recognition	Zhipeng Zhang, Xinya Peng, Xiaohang Yuan, Wenting Ma, Qingchen Liu, Yingteng He, Xingang Chai, Jie Wei, Jinman Lin, 2023	Adaptive Thresholding	Accurate Color-Strip Segmentation	96.80 %	Quality Assessment, Error Correction	Real-world application validation
[12]	Bangla and Nagri Language Detection and Classification Using Deep Learning Methods on Text Document Image	Md. Abu Naser Mojumder, Ashraful Islam, Tanzidul Islam, Showman Das, Zakaria Chowdhury, Ashfakur Rahman, 2023	Image Normalization, Noise Removal	Feature Extraction for Language Detection	95.89 %	Data Labelling	Extraction of Nagri to Bangla word
[13]	An Enhanced Deep Learning Model for Handwritten Tamil Character Identification	Babitha Lincy R, Jency Rubia J, Sherin Shibi C, Kavitha M, 2023	Image Enhancement, Slant Removal	CNN for Handwritten Character Identification	83%	Comparison to SVM	Overlapped characters and hybridization techniques
[14]	Text Extraction from Stone Inscriptions and Translating to Modern Language	Meher Pranav Kurapati, Hemanth Chowdary Vallabhaneni, Lohith Reddy Burramukku, Harshith Reddy Soda, Murali Mohan Vutukuru, Satish Thatavarthi, 2023	Image Restoration, Geometric Transformation	Various Techniques Including OCR	90%	Character Segmentation, Image Translation	Method improvement and use in other languages

[15]	A Novel Approach for Assisting Blind People Using a Smart Wearable Device	Shyam Peraka, SK Irfan Ali, Reddy Sudheer, Pilli Praveen Kumar, Goutham Kondala, Dimple Samal, 2023	Sensor Data Filtering, Data Fusion	Text Recognition, Feature Extraction from Camera and Motion Data	92%	Real-time Obstacle Detection	Advanced UI and interaction
[16]	Analysing discrete self supervised speech Representation for spoken language modelling	Amitay Sicherman and Yossi Adi, 2023	Discrete Speech Representation Analysis	Cosine Similarity Metric and DNN Prediction	80%	ABX Evaluation	Further exploration of discrete speech representation
[17]	Document-Level Neural Machine Translation With Recurrent Context States	YUE ZHAO 1 AND HUI LIU2,2023	experimentation with a method Transformer-based system.	comparison of computational and spatial complexities of two widely used document-level methods	81.40 %	neural machine translation	comparable or even better performance with strong context-aware models
[18]	OCR using CRNN: A Deep Learning Approach for Text Recognition	Aditya Yadav, Shauryan Singh, Muzzamil Siddique, Nileshkumar Mehta, Archana Kotangale,2023	Creation of a grayscale version	Connectionist Temporal Classification, these label distributions are transformed into label sequences	92.2%	Use of convolutional layers and two LSTM layers	Capability of using a CRNN architecture, which combines CNN and RNN.
[19]	Two Decades of Bengali Handwritten Digit Recognition: A Survey	A. B. M. Ashikur Rahman, Md. Bakhtiar Hasan, Sabbir Ahmed, Tasnim Ahmed, Md. Hamjajul Ashmafee, Mohammad Ridwan Kabir, Md. Hasanul Kabir, 2022	Various Preprocessing Methods	Four Categories of Descriptors for Recognition	98.80 %	Multilingual Digit Recognition, Handwritten Mathematical Expression Evaluation	Better recognition of offline Bengali handwritten digits

[20]	End to End Invoice Processing Application Based on Key Fields Extraction	Halil Arslan, 2022	Invoice Structure Analysis, Image Resizing	Table Cell Image Extraction	98.19 %	Table Line Detection	Tables for Different Companies and Documents
[21]	PHTI: Pashto Handwritten Text Imagebase for Deep Learning Applications	Ibrar Hussain, Riaz Ahmad, Siraj Muhammad, Khalil Ullah, Habib Shah, Abdallah Namoun, 2022	Skew and Line Segmentation	Gender and Age Classification	89.40 %	Character Error Rate and Word Error Rate	More data and features for recognition
[22]	Cursive Text Recognition in Natural Scene Images Using Deep Convolutional Recurrent Neural Network	Asghar Ali Chandio, Md. Asikuzzaman, Mark R. Pickering, Mehwish Leghari, 2022	Segmentation -Free Method	Feature Extraction and Sequence Transformation	95.75 %	Use of Deeper CNN Architectures	Natural scene Urdu text recognition model development

III. ALGORITHMIC SURVEY

Sr No	Paper Title	Algorithm Used	Time Complexity	Space Complexity	Accuracy (%)	Advantages/Disadvantages
1	Applying Gradient Descent in Convolutional Neural Networks	CNN	Varies	Varies	90% - 99%	Advantages: Excellent for image recognition, Disadvantages: Requires large datasets and computation
2	Exploring the Use Cases of Standard Vector Machines	SVM	O(N^2) (for training), O(N) (for classification)	O(N^2) (for training), O(N) (for classification)	85% - 95% (with appropriate kernel)	Advantages: Effective for both linear and non-linear data, Disadvantages: Can be computationally expensive
3	Text Identification from Handwritten Data using Bi-LSTM and CNN with FastAI	LSTM	O(N * M^2) (for training), O(N * M) (for inference)	O(N * M)	80% - 99% (in sequence prediction tasks)	Advantages: Good for sequential data, Disadvantages: Sensitive to hyperparameters
4	An optical character recognition algorithm for documents with cascading structures	LBP	O(N)	O(2^P) (for P patterns)	60% - 90%	Advantages: Simple and computationally efficient for texture analysis, Disadvantages: Limited to local patterns
5	Performance of Convolutional Recurrent Networks for handwritten Text Recognition	CRNN	O(N)	O(N)	High (for encryption)	Advantages: Combines both the elements of CNN and RNN, Disadvantages: Can be computationally intensive
6	Understanding Convolutional Neural Networks with A Mathematical Model	KNN	O(N * M) (for training), O(N) (for classification)	O(N * M) (for training), O(1) (for classification)	Varies (depends on data and k value)	Advantages: Simple, easy to understand, Disadvantages: Sensitive to data distribution and choice of k
7	Covid Analysis using Recurrent Neural Network	RNN	O(N * M^2) (for training), O(N * M) (for inference)	O(N * M)	80% - 99% (in sequence prediction tasks)	Advantages: Suitable for sequential data, Disadvantages: Vulnerable to vanishing gradient problem
8	Image Enhancement and Detection of Courier Slips Based on Affine Transform	Canny Edge Detection	O(N) (for convolution)	O(N)	N/a	Advantages: Detects edges accurately, low false positives. Disadvantages: Sensitive to noise, parameter tuning required.

9	Improving Classification of Scanned Document Images using a Novel Combination of Pre-Processing Techniques	Gaussian Blur	O(N) (for convolution)	O(N)	N/a	Advantages: Smooths images, reduces noise. Disadvantages: Can blur fine details.
10	An end to end neural network for image to audio transformation	ReLU (Rectified Linear Unit)	O(N)	O(N)	N/a	Advantages: Simple and computationally efficient activation function. Eliminates vanishing gradient problem. Disadvantages: Can suffer from dying ReLU problem, not suitable for all types of data.

IV. FIELD SURVEY

Date of Visit: 21 October 2023

Location: Poona School and Home for the Blind Girls, Survey No.36/3, Gandhi Bhavan Road, Kothrud, Pune, Maharashtra 411038

The objective of our field survey was to visit Poona School and Home for the Blind girls to learn more about the difficulties and problems faced by the blind to help us develop a device for their aid. Our aim is to understand the unique requirements and barriers that blind girls face when going to school and receiving education and support. When we visited the school, we observed and conducted interviews. To identify specific barriers to accessing and participating in education, we conducted interviews with educators and learners. We asked them about their experiences using public transportation and the difficulties they encountered in getting to school. We asked them whether they had access to technology and assistive devices to help them perform daily tasks. We evaluated the availability of Braille books, audio materials, and digital resources for learning and entertainment. We asked about Internet access and information collection technology.

Our fieldwork found that limited access to Braille materials, inadequate use of assistive technologies, and the need for additional support staff were major problems. Navigating unfamiliar environments proved difficult due to obstacles and lack of tactile cues, and students had difficulty using public transport independently. In terms of daily life, a significant number of students and teachers demonstrated impressive levels of independence, although there were reports of limited access to devices and assistive technology. Access to information, an important aspect of modern education and everyday life, has been hampered by the limited number of Braille books and accessible digital resources, as well as the Compatibility issue with on-screen digital reader for Internet access. Addressing these findings will be important to improve the quality of life and overall educational experience of people with visual impairments.

The recommendations we received were to develop functionalities in our device that would:

1. Read Currencies:
   An essential feature, as visually impaired individuals often struggle to identify different denominations of currency notes. The ability to recognize currency values independently would enhance their financial independence.

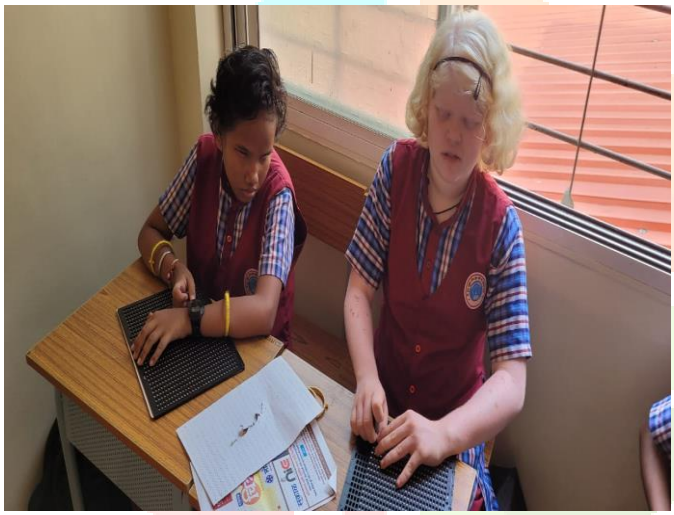
2. Read colors:
   This feature would help visually impaired individuals identify and differentiate between various objects based on color, improving their overall perception of the world.

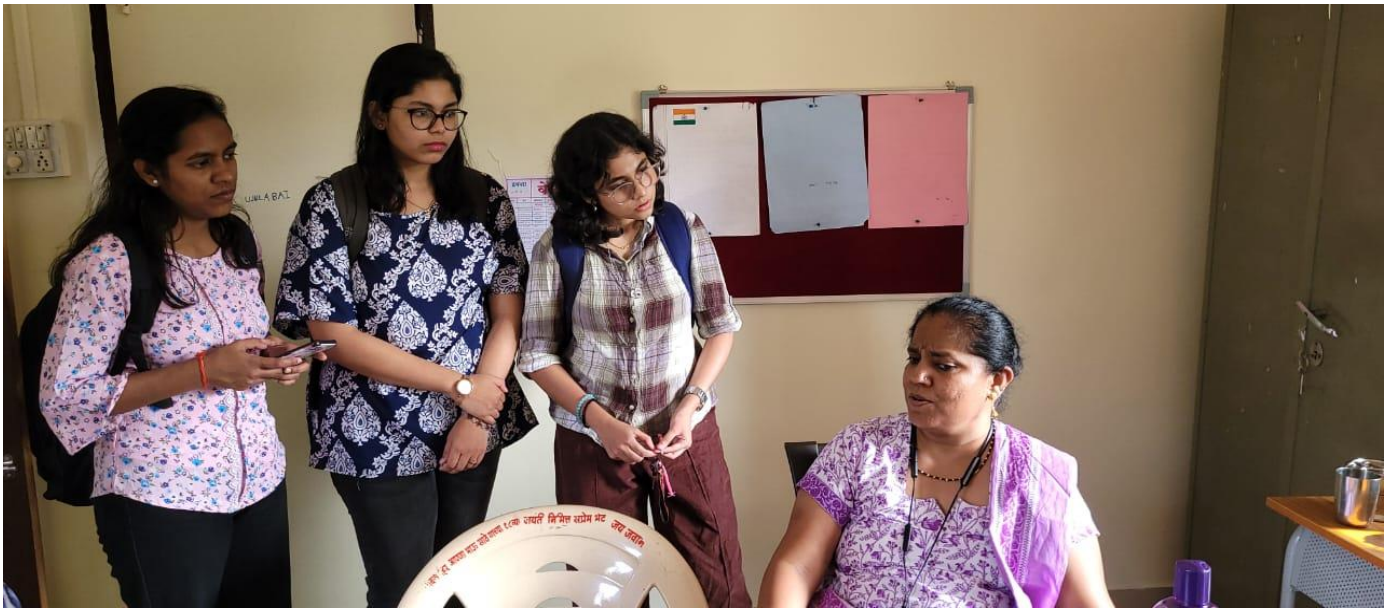
3.Provide Location Tracking:
   Location tracking is crucial for independence and safety. It can help individuals navigate through unfamiliar environments, find specific places, and ensure their security.

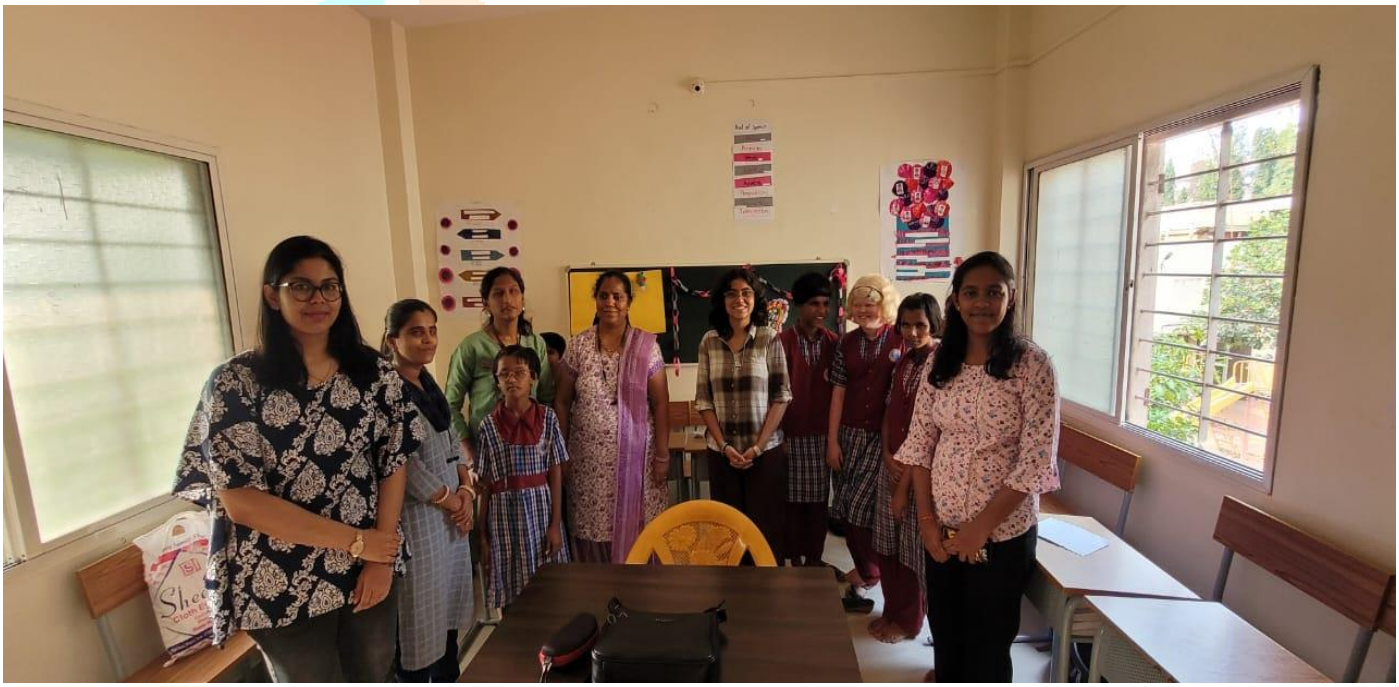
4.Read Bus Numbers and Shop Names Quickly:
   Public transportation and shopping are integral aspects of daily life. Being able to quickly identify bus numbers and shop names can significantly enhance their autonomy and ease of travel and shopping.

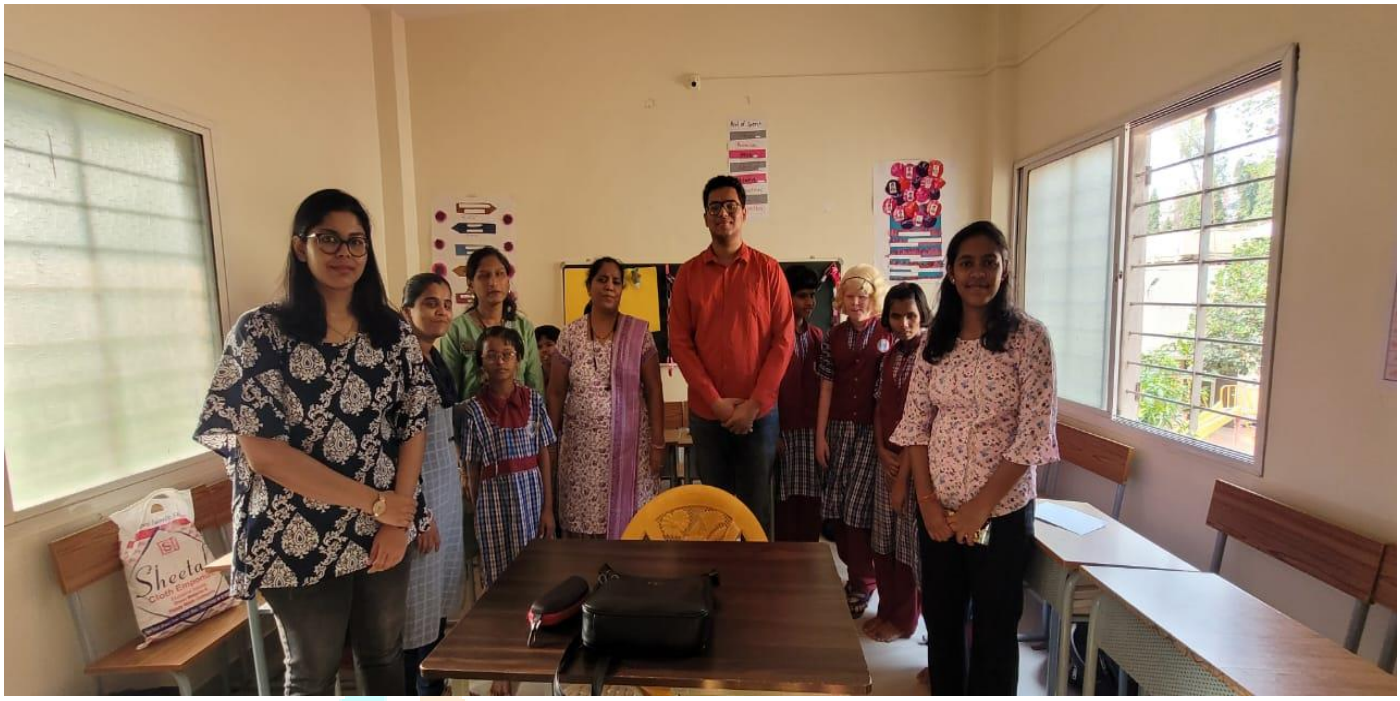
5.Detect and Inform about Stairs and Potholes:
   Identifying obstacles like stairs and potholes is essential for their safety while moving about. This feature can prevent accidents and injuries.

6.On/Off Voice Function for Convenience:
   Providing the ability to turn the voice function on and off at their discretion offers flexibility and convenience, especially in situations where voice guidance may not be required









V. CONCLUSION

The integration of OCR into smart sunglasses marks a significant stride towards progress in the field of wearable technology but also underscores the potential to enhance global connectivity and understanding. The core functionality lies in the ability to recognize and interpret text in real-time through embedded cameras and advanced image processing algorithms. By capturing and processing textual information from the user's immediate surroundings, enabling instant language recognition.

The literature survey presented research papers, each addressing different aspects of OCR-based technologies, such as language recognition, image preprocessing, and feature extraction. These studies showcased the remarkable progress with a focus on improving accuracy and expanding its applicability to diverse languages and scripts.
The algorithm survey offered insights into the computational techniques used in OCR systems. It highlighted the advantages and disadvantages of various algorithms, from [3]Convolutional Neural Networks (CNN) for image recognition to Long Short-Term Memory (LSTM) for sequence prediction. These algorithms play a crucial role in OCR's ability to accurately recognize text and provide real-time language translations.

The field survey, conducted at Poona School and Home for the Blind Girls, brought to light the challenges faced by visually impaired individuals in their daily lives, particularly in education, transportation, and access to information. The recommendations derived from this survey underscore the need for OCR-based devices to incorporate features such as currency and colour recognition, location tracking, quick identification of bus numbers and shop names, and obstacle detection.

Beyond the technological marvel, it holds the promise of fostering greater understanding and connection among individuals from diverse linguistic backgrounds. As these devices continue to evolve, their impact on global communication and accessibility is likely to shape a future where language is no longer a barrier but a bridge to shared understanding and collaboration.

REFERENCES

[1] C. H. Wen and L. L. Tze, "Optical Character Recognition with Word Prediction Feature using First Order Hidden Markov For The Blind," 2021 IEEE 12th Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, 2021, pp. 81-86, doi: 10.1109/ICSGRC53186.2021.9515248.

[2] S. Agrawal and N. Agrawal, "Recogniition and Speech Conversion of Devnagri Script using CNN," 2023 2nd International Conference for Innovation in Technology (INOCON), Bangalore, India, 2023, pp. 1-4, doi: 10.1109/INOCON57975.2023.10101034.

[3] K. Rao N.S., S. Kumar S., S. Santhosh, S. Rao and K. Kaushik P., "Smart Reader for Visually Impaired," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 846-850, doi: 10.1109/I-SMAC49090.2020.9243598.

[4] M. Raja, J. Deny, N. P and M. V, "A Peculiar Reading System for Blind People using OCR Technology," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1601-1605, doi: 10.1109/ICEARS56392.2023.10085372.

[5] B. Vidhale, G. Khekare, C. Dhule, P. Chandankhede, A. Titarmare and M. Tayade, "Multilingual Text & Handwritten Digit Recognition and Conversion of Regional languages into Universal Language Using Neural Networks," 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021, pp. 1-5, doi: 10.1109/I2CT51068.2021.9418106.

[6] P. Kapgate, S. Tidke, R. Fender, S. Rathore, S. Ghodmare and L. Ritla, "Raspberry Pi Based Book Reader For Visual Impaired People," 2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), Nagpur, India, 2023, pp. 1-6, doi: 10.1109/ICETET-SIP58143.2023.10151605.

[7] A. A, B. D. J, S. P. Mary and D. U. Nandini, "Portable Camera based Identification System for Visually Impaired People," 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2023, pp. 1444-1450, doi: 10.1109/ICOEI56765.2023.10126008.

[8] M. A. Zaryab and C. R. Ng, "Optical Character Recognition for Medical Records Digitization with Deep Learning," 2023 IEEE International Conference on Image Processing (ICIP), Kuala Lumpur, Malaysia, 2023, pp. 3260-3263, doi: 10.1109/ICIP49359.2023.10222038.

[9] L. Chen et al., "OCR Text Region Detection Based on Chinese Character Features in Natural Scenes," 2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi'an, China, 2023, pp. 1538-1541, doi: 10.1109/ICSP58490.2023.10248635.

[10] Z. Zhang et al., "Multi-connected Tiny Object Detection for Color-Strip Recognition," 2023 4th Information Communication Technologies Conference (ICTC), Nanjing, China, 2023, pp. 417-422, doi: 10.1109/ICTC57116.2023.10154638.

[11] M. Abu Naser Mojumder, A. Islam, T. Islam, S. Das, Z. Chowdhury and A. Rahman, "Bangla and Nagri Language Detection and Classification Using Deep Learning Methods on Text Document Image," 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), Bangalore, India, 2023, pp. 480-485, doi: 10.1109/ICAECIS58353.2023.10170588.

[12] R. Babitha Lincy, J. Jency Rubia, C. Sherin Shibi and M. Kavitha, "An Enhanced Deep Learning Model for Handwritten Tamil Character Identification," 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2023, pp. 1518-1523, doi: 10.1109/ICSSIT55814.2023.10060920.

[13] M. P. Kurapati, H. C. Vallabhaneni, L. R. Burramukku, H. Reddy Soda, M. M. Vutukuru and S. Thatavarthi, "A Review: Text Extraction from Stone Inscriptions and Translating to Modern Language," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1515-1519, doi: 10.1109/ICEARS56392.2023.10085645.

[14] S. Peraka, S. I. Ali, R. Sudheer, P. P. Kumar, G. Kondala and D. Samal, "A Novel Approach for Assisting Blind People Using a Smart Wearable Device," 2023 36th International Conference on VLSI Design and 2023 22nd International Conference on Embedded Systems (VLSID), Hyderabad, India, 2023, pp. 383-388, doi: 10.1109/VLSID57277.2023.00082.

[15] A. Sicherman and Y. Adi, "Analysing Discrete Self Supervised Speech Representation For Spoken Language Modeling," ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10097097.

[16] Y. Zhao and H. Liu, "Document-Level Neural Machine Translation With Recurrent Context States," in IEEE Access, vol. 11, pp. 27519-27526, 2023, doi: 10.1109/ACCESS.2023.3247508.

[17] A. Yadav, S. Singh, M. Siddique, N. Mehta and A. Kotangale, "OCR using CRNN: A Deep Learning Approach for Text Recognition," 2023 4th International Conference for Emerging Technology (INCET), Belgaum, India, 2023, pp. 1-6, doi: 10.1109/INCET57972.2023.10170436.

[18] A. B. M. Ashikur Rahman et al., "Two Decades of Bengali Handwritten Digit Recognition: A Survey," in IEEE Access, vol. 10, pp. 92597-92632, 2022, doi: 10.1109/ACCESS. 2022.3202893.

[19] H. Arslan, "End to End Invoice Processing Application Based on Key Fields Extraction," in IEEE Access, vol. 10, pp. 78398-78413, 2022, doi: 10.1109/ACCESS.2022.3192828.

[20] I. Hussain, R. Ahmad, S. Muhammad, K. Ullah, H. Shah and A. Namoun, "PHTI: Pashto Handwritten Text Imagebase for Deep Learning Applications," in IEEE Access, vol. 10, pp. 113149-113157, 2022, doi: 10.1109/ACCESS.2022.3216881.

[21] A. A. Chandio, M. Asikuzzaman, M. R. Pickering and M. Leghari, "Cursive Text Recognition in Natural Scene Images Using Deep Convolutional Recurrent Neural Network," in IEEE Access, vol. 10, pp. 10062-10078, 2022, doi: 10.1109/ACCESS.2022.3144844.

[22] Nan Cui[1] , "Applying Gradient Descent in Convolutional Neural Networks",in IOP Conf. Series: Journal of Physics: Conf. Series 1004 (2018) 012027 , doi :10.1088/1742-6596/1004/1/012027

[23] R. S. Srichandra, P. S. Rahul, M. S. Govind and V. Battula, "Performance of Convolutional Recurrent Networks for Handwritten Text Recognition," 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2023, pp. 1207-1210.

[24] X. Wang, F. Xu, Y. Ye and Y. Xu, "Image Enhancement and Detection of Courier Slips Based on Affine Transform," 2023 9th International Conference on Control, Automation and Robotics (ICCAR), Beijing, China, 2023, pp. 350-354, doi: 10.1109/ICCAR57134.2023.10151736.

[25] A. D. Mahajan, S. Karuppasamy and S. Lakshminarayanan, "Improving Classification of Scanned Document Images using a Novel Combination of Pre-Processing Techniques," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 1109-1114, doi: 10.1109/ICAAIC56838.2023.10140911.

[26] Manoj Kumar Padhi, Sahana D S, Bindu Madhavi K, Bechoo Lal, Karunakara Rai B, "Covid Analysis using Recurrent Neural Network", 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), DOI: 10.1109/ICICACS57338.2023.10100129

[27] L. Chen, M. Deisher and M. Georges, "An End-to-End Neural Network for Image-to-Audio Transformation," ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10096121.