



e-ISSN: 2278-8875

p-ISSN: 2320-3765

International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering

Volume 11, Issue 3, March 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.282

☎ 9940 572 462

☑ 6381 907 438

✉ ijareeie@gmail.com

@ www.ijareeie.com



IOT Based Optical Character Recognition Modelled Character Recognition System for Visually Impaired Persons

K. Suresh Kumar^{*1}, T. Ananth Kumar², S. Sundaresan³

Associate Professor, Electronics and Communication Engineering, IFET College of Engineering, India

Assistant Professor, Computer Science and Engineering, IFET College of Engineering, India

Research Scholar, Electronics and Communication Engineering, National Institute of Technology, Puducherry, India

ABSTRACT: Many people around the world have physical disabilities. People who fall into this category need help from a third person because they have more problems in their daily lives. This is why they need help from someone else. Reading books, doing other things, and crossing the street are a few of the things they have to deal with. The route number is shown only on the bus number plate. This can be a problem at bus stops that aren't used very often. Even though the bus's location has been found, the user does not know where he or she is supposed to go. Use the skills of people who cannot see or see very little to make new systems and apps for systems used by the general public already. OCR (Optical Character Recognition) is used to make a book-reading system for people who are blind. In this way, the biggest concern with self-reading has been rectified.

KEYWORDS: *IoT; Blind persons; OCR; Device;*

I. INTRODUCTION

By 2050, there will be 115 million blind people worldwide, up from 36 million today because the world's population will grow. This embedded device is only for blind or visually impaired people[1]. Blind people have a hard time living, but technology can help them. Blind people struggle with technology as well. People without eyes cannot read the information on a web page. It takes practice to use screen reading software to search the internet. Because screen reading software reads website content, this work focuses on developing a text-to-speech app for the blind. The Raspberry Pi is the star of this project[2]. It has a built-in camera that can scan any document and convert it to digital text using optical character recognition. To create the voice, we use a system that can synthesize digital text. It is part of this project to look at the text and do OCR (Optical Character Recognition) (OCR). HMM is used to match input to learned data.

After removing noise from images, low-level feature extraction is used to look for lines, regions, and possibly textures. In this case, think of them as single objects, like cars or boxes on a conveyor belt [3]. One issue with AI is that things can appear very different depending on the angle or light. Also, difficult to determine which features belong to an object and which belong to the background, shadows or other non-objects[4]. The human visual system does this automatically, but a computer requires skilled programming and a lot of processing power to match human performance. Photographic prints, slides, TV screens, and movie screens are all images. A camera or a computer can alter an image.

To digitally process an image, convert it to a set of numbers that a computer can manipulate. A pixel is a number that shows how bright an image is at a specific location. Larger images are becoming more common[5]. A typical digital image is 512x512 or roughly 250,000 pixels. After digitizing images, three basic things can be done with them. A point operation modifies one pixel in the input image—the pixel value in the output image changes as a result. For local operations, the value of a pixel in the output image is determined by several pixels in the input image. When using global operations, all pixels in the input image are used to create the output image's value[6].

These actions, alone or together, improve, restore, or compress the image. An image is enhanced when it is changed to make the content more visible. This can also improve the image's appeal. A 3 x 3-pixel window can be used to improve a noisy image[7]. In this case, each pixel in the noisy image is recorded along with its eight neighbors. The nine numbers are then sorted by size, and the median is used for each pixel in the new image. The filtered image is made as to the 3 x 3 window moves across the noisy image[8]. The user shall use contrast manipulation to improve the image. A point operation is one where each pixel in the new image has the same value as in the old image. Changing



the brightness and contrast on a TV or changing the exposure and development time on a print usually alters the contrast. Another point operation is to colorize a black-and-white image by assigning random colors to the grey levels[9]. This method is used in thermographs, which show heat. In this case, hotter objects (high pixel values) have one color and cooler ones (low pixel values). Intermediate values get other colors. By 2050, there will be 115 million blind people worldwide, up from 36 million today because the world's population will grow. This embedded device is only for blind or visually impaired people. Blind people have a hard time living, but technology can help them. Blind people struggle with technology as well. People without eyes cannot read the information on a web page. It takes practice to use screen reading software to search the internet[10]. Because screen reading software reads website content, this work focuses on developing a text-to-speech app for the blind. The Raspberry Pi is the star of this project. It has a built-in camera that can scan any document and convert it to digital text using optical character recognition. To create the voice, we use a system that can synthesize digital text. It is part of this project to look at the text and do OCR (Optical Character Recognition) (OCR). HMM is used to match input to learned data. After removing noise from images, low-level feature extraction is used to look for lines, regions, and possibly textures[11]. In this case, think of them as single objects, like cars or boxes on a conveyor belt. One issue with AI is that things can appear very different depending on the angle or light. Also difficult to determine which features belong to an object and which belong to the background, shadows or other non-objects. The human visual system does this automatically, but a computer requires skilled programming and a lot of processing power to match human performance. Photographic prints, slides, TV screens, and movie screens are all images. A camera or a computer can alter an image.

To digitally process an image, convert it to a set of numbers that a computer can manipulate. A pixel is a number that shows how bright an image is at a specific location. Larger images are becoming more common. A typical digital image is 512512 or roughly 250,000 pixels. After digitizing images, three basic things can be done with them. A point operation modifies one pixel in the input image—the pixel value in the output image changes as a result. For local operations, the value of a pixel in the output image is determined by several pixels in the input image. When using global operations, all pixels in the input image are used to create the output image's value[12].

These actions, alone or together, improve, restore, or compress the image. An image is enhanced when it is changed to make the content more visible. This can also improve the image's appeal. A 3 x 3-pixel window can be used to improve a noisy image. In this case, each pixel in the noisy image is recorded along with its eight neighbors. The nine numbers are then sorted by size, and the median is used for each pixel in the new image[13]. The filtered image is made as to the 3 x 3 window moves across the noisy image. The user shall use contrast manipulation to improve the image. A point operation is one where each pixel in the new image has the same value as in the old image. Changing the brightness of a TV or the exposure or development time of a print usually alters the contrast. To add color to a black-and-white photograph, use a random selection of colors for the grayscale. In thermography, heat is shown through this technique. In this case, hotter objects (high pixel values) have one color and cooler ones (low pixel values). Intermediate values get other colors[14].

For a long time, computer vision researchers have been attempting to figure out what is in images. The appearance of a class can take on a surprising variety of different shapes and sizes. As a result, our ability to comprehend is impaired. Various factors, including clutter, scale, and perspective, can impact the appearance of even a single object. Furthermore, objects from different social strata may have striking visual similarities with one another. This is highly significant. It is necessary for models of object classes to deal with a wide range of classes to distinguish between real and fake objects in crowded images[15]. Because of the large number of conflicting rules, it is difficult to distinguish between a class model and another. A major focus of this paper is on how to recognize and classify objects in images, which is a crucial component of the research. Images can be classified as belonging to an object class if they contain elements of the object class in question. Attempting to locate all instances of a particular class within a given image may yield results in some instances[15]. Objects can be classified solely on the basis of their edges, which is the primary goal of this paper. Compared to other learning methods, this method is distinguished by the use of simple line segments and ellipses as primitives. A model that travels in the same direction as the line is available, whereas a model that travels in the opposite direction is unavailable. Despite the fact that it is one of the simplest circular shapes, it can display complex curves. Despite their primitive appearance, these forms contain a plethora of beneficial elements. To put it another way, they are more versatile than edge-based descriptors in that they can be used to draw both abstract and real-world conclusions. There is no difference in storage space between a line with four parameters and an ellipse with five parameters[16].

Primitive matching compares all of the edge pixels simultaneously, rather than comparing each one at a time like traditional matching. When scaling, geometric properties can be normalized, which makes the process of scaling easier. Because contour fragments are not scale-invariant, the image must first be resized or rescaled before they can be extracted. When the pixels are pulled apart, this can result in aliasing or blurry edges on the edges. Because lines and ellipses are generic shapes and structures, they can be used to represent complex shapes and structures in images. When



these primitives are combined, they become sufficiently distinct to be understood. Initially, shape token pairs are used to construct the first layer of combinations, and the second layer is constructed from a learned number of shape tokens. The number of shape tokens is not fixed, but they must be able to change to match the type of object they represent in the game. Generally speaking, shape tokens are used less frequently in games where the shapes are more complex. Then, it is possible to create complex discriminative combinations of objects for a given class of objects. When defining a shape constraint, it is essential to consider the appearance of the shape token being used. Geometric constraints are the most important considerations when it comes to real-world appearances (configurations). People who are in shape-token relationships impose structural constraints on the system, which negatively impacts its performance (e.g. XOR relationships).

II. LITERATURE SURVEY

Microelectromechanical systems (MEMS) sensors are improving human-computer interaction. They are cheap because they have no embedded hardware and can write. Today's inertial pen character recognition methods use low-level inertial signal features. The authors present an inertial sensor-based Persian/Arabic handwriting recognition system[17]. The inertial pen's motion trajectory is first reconstructed to derive position signals. Use these signals to get high-level geometric features from data. A new metric learning technique improves character classification accuracy. A genetic programming algorithm calculates each character's characteristic function. They created a metric kernel classifying all Persian and Arabic handwriting recognition characters that outperformed current work[18].

This application is aimed at disabled people who cannot use mobile phones or other communication devices. That is, voice commands can be used for everything. We're talking about a voice-based messaging system. The app converts text to speech. Voice-to-Text technology on Android makes it simple to communicate with loved ones via social media[19]. By name, you would expect it to be about SMS. Standard protocols allow text message transmission between two or more mobile phones. SMS text messaging is the most popular data application for mobile phone users, with 2.4 billion active users (74 percent). The app will recognize and convert speech. ' Android does support underrepresented groups often overlooked by the general public. Disabled people have a more challenging time navigating society than ordinary people. But they will always enjoy their daily routines. This app was created for people with disabilities who struggle with communication. This app's main function is to send and receive SMS. You can now read your voicemail in text[20].

A mobile camera detects an object and outputs audio. Using features to locate an object. Android, Google's mobile operating system, does not currently support voice commands. See what your phone can do with the Services[21]. Thus, the goal is to make Android accessible to the blind. Blind people want to use voice commands to make calls and read text messages on Android phones. Object detection is a rapidly expanding field. Image processing can detect and track moving objects in industrial and social contexts[22]. This work will provide object detection and audio/vocal information to the blind. We use the camera on our phones to locate objects and give voice commands. Before accessing object data, users must be trained. Then On uses feature extraction to identify objects in the camera view[23]. We use the object's placement angle as a compass.[28]This work is developing a Tamil character recognition system. A handwritten character is made up of strokes. In this method, a stroke is a random string of shape features[24][25]. Unknown strokes can be identified and tracked using flexible string matching. All of its strokes must be identified to identify a character as a whole. A finite-state automaton controls a character's progress. The development of similar systems for other Indian scripts is outlined.[26,27]

III. PROPOSED SYSTEM

If the user uses this advanced character based recognition system, user will have no trouble distinguishing between English and non-English content. The characters are recognized and segmented with the help of optical character recognition (OCR) and a hidden Markov model. As a result, a well-known persona or phrase is employed in this context. This is accomplished by converting the RGB image into binary data (see below). During the first two stages of the segmentation process, the horizontal projection profile of a document is used to identify and divide a document's lines into segments based on the lines' orientation. When it comes to sorting words and characters in documents, vertical projection profiles are used similarly as horizontal projection profiles. When comparing the segmented characters to previously defined characters and words, the hidden markov method is used.

Splitting an image into two glyphs is the first step in optical character recognition (OCR). Two separate phases are necessary for our method, each with slightly varying requirements. Segmented glyphs for classic manual literature are shown to a human supervisor during training. As soon as we have been trained, we want to recognize a new image. First, we need to identify each glyph before extracting features and categorizing them. I first try to divide the image



into lines for the second step. Binary thresholds are used to process our images. Figure 1 shows the average intensities across each row, which tells us where the gaps between the lines are in this case. Figure 1: (mean intensity close to 1). As a result, we're able to spot image gaps and separate the resulting pieces. This rudimentary implementation fails when a paragraph's final lines are only a few words long. The calculation of line gaps from mean row intensities is implemented as a separate function that can be improved without affecting the rest of the procedure. Word split lines can also be made using a similar technique. A good preprocessing step is line segmentation. When a matrix with 0 and 1 entries is passed to the segmentation code, it returns a matrix with 0, 1, 2, 3, N entries, where N is the number of segments that were found. The i th segment is represented by the matrices I , $I = 2$, and N . C code is called in for this computationally intensive part of R. The segments are then extracted as binary matrices using another R function. Classifiers are trained using segmentation, as previously mentioned. For training purposes, several glyphs must be manually identified. There are a variety of ways to go about this. This is what I'm going to do from here on out. Binary matrixes with manually selected classes keep track of the glyphs that have been used for training. R binary format is used to store this data in a file named "trainingData.rda." The function `updateTrainingSet` can be used to add images to this list. A glyph is shown to the user before they are asked to enter their class. The class name and corresponding matrix are added to the list. The glyph is ignored if no class is specified.

It generates binary matrices that are unsuitable for use in neural networks. Because each glyph has its own set of characteristics, we must first determine what those characteristics are. Neither of these options is the most desirable. It is easier to create an image by putting the binary glyph matrix into a set of points rather than the other way around. In this example, pixels in the foreground of the image with at least one background neighbor (one) are considered out of focus (1). A binary matrix can be found on the left-hand side of the screen. Each of the 'boundary' points is depicted separately on the right, next to each other. The local limit is defined as follows: Each point has its own set of steps, which are as follows: Then, to determine the direction of the boundary, draw a straight line through a few nearby points using a straight line. Then, after rotating the image, you can fit a quadratic regression line to the points on the image that you already know. The points on this quadratic curve are in perfect alignment with one another (using a predetermined threshold). To make the quadratic curve fit, use the information from the previous points. The term "close points" appears twice in this new update, as you can see. According to the data collected, the curves discovered on either side of the line are consistent with those discovered on the other. It is acceptable as long as it does not perform flawlessly in every situation. We are interested in finding the curves that fit the most significant number of points on the boundary. This type of curve will have many points, and some of them will be able to identify the shape of the curve correctly. At the end of the process, we look at each point and rank them according to their proximity to the best fit curve. It is necessary to record the rotational angle and quadratic coefficient of the curve that provides the best fit (the linear coefficient is close to 0 because of the rotation). In order to find a glyph, we search for the second and third best curves. Consequently, we pay no attention to any points in the vicinity of the first curve.

1.1. Experimental Setup

The purpose of this research is to investigate the manner in which the English alphabet is used on public signs. It was decided to use the built-in camera on the phone to photograph common public signs, which included things such as wet floors on the sidewalks, bus stops that were closed or open, doors that were closed or open, an exit that was closed or open, a taxi that was stopped, restrooms, and road signs, among other things. A group of participants who were blindfolded took photographs, which were then uploaded to an online database, where they were compared to photographs taken by a group of participants who were not blindfolded. 10 females and 10 males who were all normally sighted were recruited to take part in the tests in order to carry out the research.



Table 1. Recognition rate of the identified samples

No.of Persons	No.of Samples Taken	Identified Samples	Recognition Rate(%)
2	12	7	58.33
6	52	47	90.38
8	48	43	89.58
10	62	57	91.94
12	65	60	92.31
14	50	45	90.00
16	10	5	50.00
20	85	80	94.12

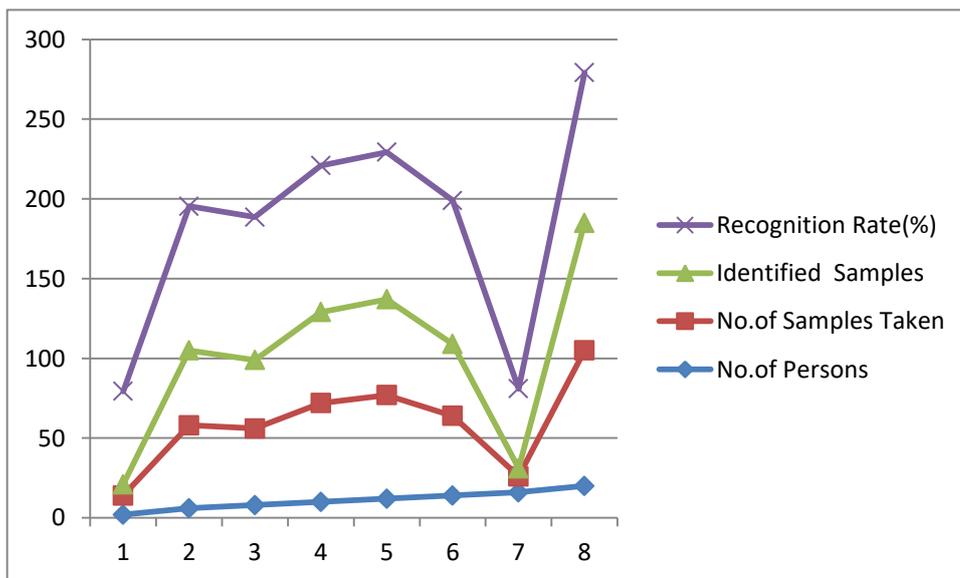


Figure 1. Analysis of recognition rate according to samples

Table 1. and Figure 1. shows the detailed analysis graph of recognition rate according to samples.

IV. CONCLUSION

For example, when microelectromechanical systems come up with sensors for inertial pens, they will become more and more popular. These writing tools are cheap because they do not need special hardware to work. For the most part, inertial pen character recognition algorithms use low-level inertial signal characteristics to figure out what letters are on the screen. At first, a pen that could read Persian and Arabic handwriting characters was shown in this part of the world. An inertial navigation system's principles are used to determine how the inertial pen moves. Then, this process starts. When the position signals are looked at again, more geometry is found. After that, a new metric learning method is used to help us better understand how to classify people. In the end, a genetic programming algorithm is used to figure out what each character's unique job is. A good handwriting recognition tool does not do as well as this one.

REFERENCES

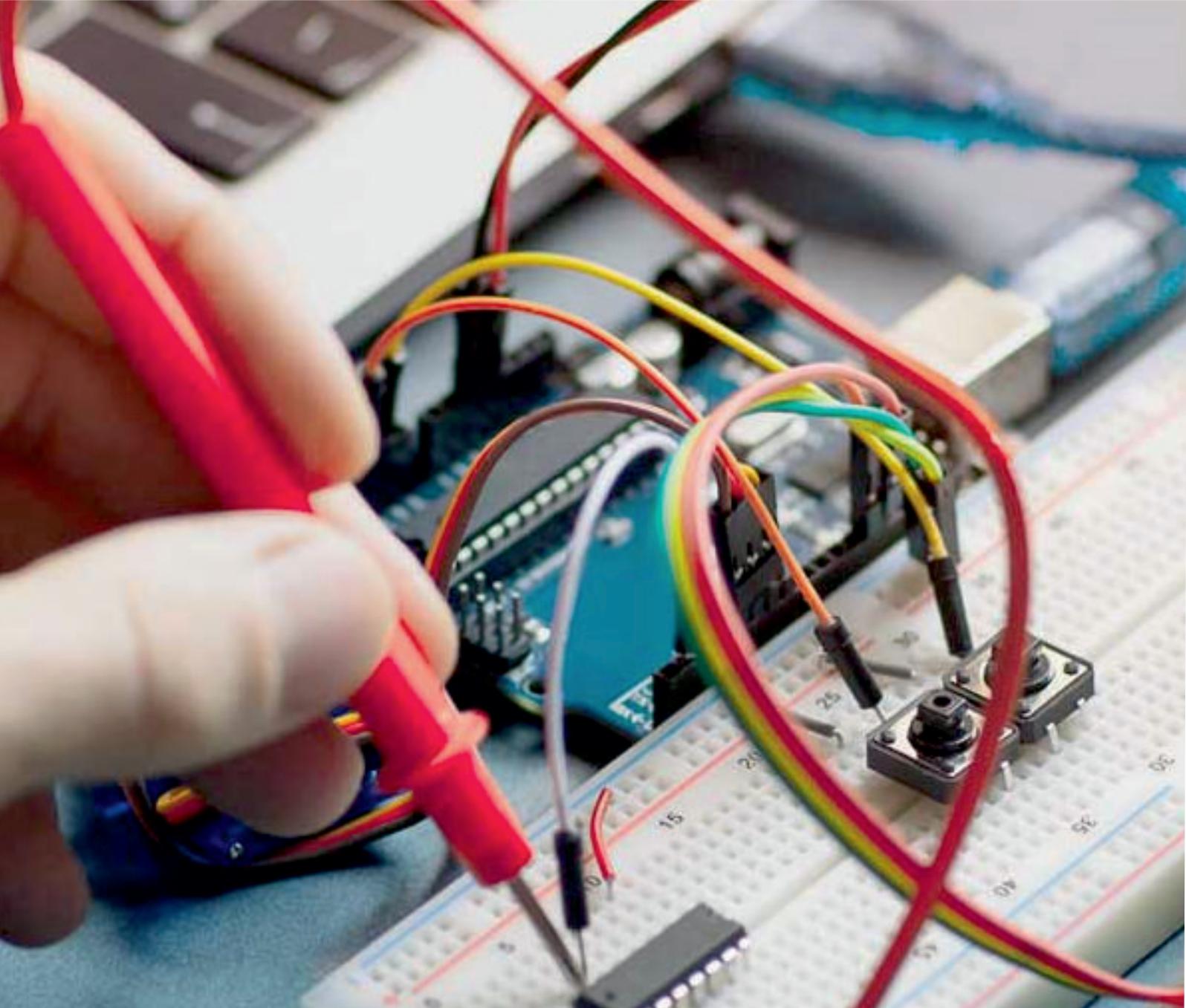
1. Van Hoai, D. P., Duong, H. T., & Hoang, V. T. (2021). Text recognition for Vietnamese identity card based on deep features network. *International Journal on Document Analysis and Recognition (IJDAR)*, 1-9.



2. Kumar, S. A., Kumar, H., Dutt, V., & Dixit, P. (2020). Deep Analysis of COVID-19 Pandemic using Machine Learning Techniques. *Global Journal on Innovation, Opportunities and Challenges in Applied Artificial Intelligence and Machine Learning* [ISSN: 2581-5156 (online)], 4(2).
3. Kumar, TamilarasanAnanth, RajendraneRajmohan, Muthu Pavithra, Sunday AdeolaAjagbe, Rania Hodhod, and Tarek Gaber. "Automatic Face Mask Detection System in Public Transportation in Smart Cities Using IoT and Deep Learning." *Electronics* 11, no. 6 (2022): 904.
4. Kumar, S. A., Kumar, H., Dutt, V., & Swarnkar, H. (2020). Role of Machine Learning in Pattern Evaluation of COVID-19 Pandemic: A Study for Attribute Explorations and Correlations Discovery among Variables. *Global Journal on Application of Data Science and Internet of Things* [ISSN: 2581-4370 (online)], 4(2).
5. Liu, X., Carrington, P., Chen, X. A., & Pavel, A. (2021, May). What Makes Videos Accessible to Blind and Visually Impaired People?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).
6. KUMAR, S. A., KUMAR, H., DUTT, V., & SWARNKAR, H. (2019). CONTRIBUTION OF MACHINE LEARNING TECHNIQUES TO DETECT DISEASE IN-PATIENTS: A COMPREHENSIVE ANALYSIS OF CLASSIFICATION TECHNIQUES. *Global Journal on Innovation, Opportunities and Challenges in Applied Artificial Intelligence and Machine Learning* [ISSN: 2581-5156 (online)], 3(1).
7. Khan, A., & Khusro, S. (2021). An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities. *Universal Access in the Information Society*, 20(2), 265-298.
8. Chen, Z., Liu, X., Kojima, M., Huang, Q., & Arai, T. (2021). A Wearable Navigation Device for Visually Impaired People Based on the Real-Time Semantic Visual SLAM System. *Sensors*, 21(4), 1536.
9. Kumar, S. A., García-Magariño, I., Nasralla, M. M., & Nazir, S. (2021). Agent-Based Simulators for Empowering Patients in Self-Care Programs Using Mobile Agents with Machine Learning. *Mobile Information Systems*, 2021.
10. Gao, C., Zhu, Q., Wang, P., Li, H., Liu, Y., Van den Hengel, A., & Wu, Q. (2021). Structured multimodal attentions for textvqa. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
11. Meraj, T., Alosaimi, W., Alouffi, B., Rauf, H. T., Kumar, S. A., Damaševičius, R., & Alyami, H. (2021). A quantization assisted U-Net study with ICA and deep features fusion for breast cancer identification using ultrasonic data. *PeerJ Computer Science*, 7, e805.
12. Tolstov, I., Martynov, S., Farsobina, V., & Bulatov, K. (2021, January). A modification of a stopping method for text recognition in a video stream with best frame selection. In *Thirteenth International Conference on Machine Vision* (Vol. 11605, p. 116051M). International Society for Optics and Photonics.
13. Kumar, S. A., Kumar, H., Dutt, V., & Dixit, P. (2020). The Role of Machine Learning in COVID-19 in Medical Domain: A Survey. *Journal on Recent Innovation in Cloud Computing, Virtualization & Web Applications* [ISSN: 2581-544X (online)], 4(1).
14. Kumar, T. A., Selvi, S. A., Rajesh, R. S., Perumal, P. S., & Stalin, J. (2021). Limos—Live Patient Monitoring System. In *Handbook of Artificial Intelligence in Biomedical Engineering* (pp. 311-330). Apple Academic Press.
15. Kumar, S. A., Kumar, H., Dutt, V., & Swarnkar, H. (2020). COVID-19 Pandemic analysis using SVM Classifier: Machine Learning in Health Domain. *Global Journal on Application of Data Science and Internet of Things* [ISSN: 2581-4370 (online)], 4(1).
16. Kumar, S. A., Nasralla, M. M., García-Magariño, I., & Kumar, H. (2021). A machine-learning scraping tool for data fusion in the analysis of sentiments about pandemics for supporting business decisions with human-centric AI explanations. *PeerJ Computer Science*, 7, e713.
17. Mostafa, A. M., Kumar, S. A., Meraj, T., Rauf, H. T., Alnuaim, A. A., & Alkhayyal, M. A. (2022). Guava Disease Detection Using Deep Convolutional Neural Networks: A Case Study of Guava Plants. *Applied Sciences*, 12(1), 239.
18. Samuel, T. A., Pavithra, M., & Mohan, R. R. (2021). LIFI-Based Radiation-Free Monitoring and Transmission Device for Hospitals/Public Places. In *Multimedia and Sensory Input for Augmented, Mixed, and Virtual Reality* (pp. 195-205). IGI Global.
19. Kumar, S. A., Kumar, H., Dutt, V., & Soni, H. (2021, February). Self-Health Analysis with Two Step Histogram based Procedure using Machine Learning. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 794-799). IEEE.
20. Sharma, S., & Salma, M. U. (2021). Social, Medical, and Educational Applications of IoT to Assist Visually Impaired People. In *Internet of Things for Healthcare Technologies* (pp. 195-214). Springer, Singapore.
21. Kumar, S. A., Kumar, A., Dutt, V., & Agrawal, R. (2021, February). Multi Model Implementation on General Medicine Prediction with Quantum Neural Networks. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 1391-1395). IEEE.
22. Sahel, J. A., Boulanger-Scemama, E., Pagot, C., Arleo, A., Galluppi, F., Martel, J. N., ... & Roska, B. (2021). Partial recovery of visual function in a blind patient after optogenetic therapy. *Nature Medicine*, 1-7.



23. Kumar, S. A., Kumar, H., Swarna, S. R., &Dutt, V. (2020). Early Diagnosis and Prediction of Recurrent Cancer Occurrence in a Patient Using Machine Learning. *European Journal of Molecular & Clinical Medicine*, 7(7), 6785-6794.
24. Real, S., & Araujo, A. (2019). Navigation systems for the blind and visually impaired: Past work, challenges, and open problems. *Sensors*, 19(15), 3404.
25. Kumar, Abhishek, SwarnAvinash Kumar, Vishal Dutt, Ashutosh Kumar Dubey, and Vicente García-Díaz. "IoT-based ECG monitoring for arrhythmia classification using Coyote Grey Wolf optimization-based deep learning CNN classifier." *Biomedical Signal Processing and Control* 76 (2022): 103638.
26. A.Kumar, S.Kumar, V.Dutt, S.Narang, A.Dubey "A Hybrid Secured Cloud Platform Maintenance based on Improved Attributes. Based Encryption Strategies" published in regular issue in *IJIMAI*, Indexed by the Science Citation Index Expanded(Web Of Science), Universidad International de La Rioja (UNIR). ISSN 1989-1660.
27. SwarnAvinash Kumar, Harsh Kumar, Vishal Dutt, HimanshuSwarnkar, "Contribution Of Machine Learning Techniques To Detect Disease In Patients : A Comprehensive Analysis Of Classification Techniques"Vol 3 No 1 (2019): *Global Journal on Innovation, Opportunities and Challenges in AAI and Machine Learning*. ISSN 2581-5156.
28. SwarnAvinash Kumar, Kapil Chauhan, AasthaParihar, "Functionality of Classification and Regression tree in Bioinformatics"Vol 5 No 2 (2021): *Global Journal on Innovation, Opportunities and Challenges in Applied Artificial Intelligence and Machine Learning*. ISSN 2581-5156.



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.282



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering

 **9940 572 462**  **6381 907 438**  **ijareeie@gmail.com**



www.ijareeie.com

Scan to save the contact details