IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

Wearable Aid for Assisting the Blind

P. Devi[1], B. Saranya[2], B. Abinayaa[3], G. Kiruthikamani[4], N. Geethapriya[5]

Assistant Professor, Sri Ramakrishna Institute of Technology, Coimbatore, India[1, 2, 3, 4]

PG Scholar, Sri Ramakrishna Institute of Technology, Coimbatore, India[5]

Abstract: Humans around us can be linked through communication. While people with hearing or visual impairments alone can find a mode to share their opinions with others and recognize them, deaf- blind people face a much more difficult communication job. The appropriate technology can play a decisive role. A deaf-blind person is one with impaired senses of hearing and sight. In this paper, we present an approach to analysis of automatic and accurate text detection and recognition of signs for blind persons. The recognized text codes are output to blind users in speech. The proposed approach embeds multi-resolution and multi-scale edge detection, adaptive searching, colour study, and affine rectification in a hierarchical framework for sign detection, with different accents at each phase to maintain the text in customized sizes, orientations, colour distributions and backgrounds. A fast and effective clipping algorithm is planned to extract Maximally Stable Extremal Regions (MSERs) as character candidates using the policy of minimizing regularized differences. The procedure can significantly improve text detection rate and optical character recognition (OCR) accuracy. Performance of the proposed text localization is quantitatively evaluated on ICDAR-2003 and ICDAR-2011 Robust Reading Datasets.

Keywords: MSER, Multi-Resolution, Optical character recognition

I. INTRODUCTION

In the European countries, it is predictable that 10–15% of the total population is disabled and the population aged 60 years and above has a ratio at nearly 1 person in 10. The ratio indicates that in EU there are about 80 million elderly or disabled people. Given the growth in life expectancy in the world, a large part of its population will experience functional problems. The sectors such as population, government and public institutions have been promoting research in this line in the recent years.
The schemes to aid communication, mobility of elderly & disabled persons, aims to increase their life quality by allowing them more autonomous, independent lifestyle and greater chances of social interaction.

Reading is obviously essential in today's society. Printed text is everywhere in the form of paper works like reports, bank statements, restaurant menus, classroom handouts, product packages, etc. Apart from this, the optical aids and video magnifiers can help blind users and those with low vision to access forms. There are many devices that can provide good access to common hand-held objects such as product sets, and objects printed with text. The ability of people who are blind or have major visual impairments to read printed tags and product packages will enhance independent living, and foster economic and social self-sufficiency. Text in images contains valuable information and is exploited in many content-based image and video bids, such as content-based web image search, video information retrieval, and mobile based text analysis and recognition. Due to composite background, and variations of font, size, colour and orientation, text in normal scene images has to be forcefully detected before being recognized and retrieved.

Sign language is a part of the live. A sign is an entity that proposes the existence of a fact. It can be a displayed structure bearing letters or codes, used to identify some object or promote a place of business. It can also be a posted notice bearing a designation, direction, safety advisory, or command.
Automatic detection and recognition of text from natural scenes is a very difficult job. The important challenge lies in the range of text: it can vary in font, size, orientation, and position. Text can also be blurred from motion or obstructed by other stuffs. As signs exist in three-dimensional space, text on signs in images displayed on the screen can be distorted by position and shape of the object. In addition to the horizontal left-to-right positioning, other alignments include vertical, circular wrapped around another object, slanted, sometimes with the typescripts tapering, and even mixed orientations within the same text space, such as script on a T-shirt or crumpled sign. Though many commercial OCR systems work well on high quality scanned papers under a controlled situation, they have a far higher error rate for sign recognition tasks because of low quality images.

II. RELATED WORK

MSER-based methods have proven promising performance in many real-time projects. However, present MSER-based approaches still have some significant limitations, i.e., they suffer from detecting iterating modules and also insufficient text entrant construction algorithms. In this section, we will review the MSER-based methods focusing on these two problems. The main advantage of MSER-based approaches over traditional

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

methods origins in the usage of the MSERs algorithm for character extraction – the MSERs algorithm is able to detect most characters even when the image is of low quality. One of the severe but not so evident pitfalls of the MSERs algorithm is that most of the detected MSERs are in detail repeating with each other. Repeating MSERs are problematic for the later character entrants grouping algorithm, apart from the MSERs that most likely correspond to character, need to be removed before being fed to the character grouping algorithm. The MSERs snipping problem has been studied by Carlos et al. and Neumann et al. Carlos et al., proposed a MSERs snipping algorithm that contains two steps: (1) reduction of linear segments by exploiting the border energy function; and (2) hierarchical filtering with a cascade of filters. Neumann et al proposed a MSER++ based text detection method, which exploits complicated features. Later, Neumann et al presented a two-stage algorithm for Extremal Regions snipping with the exhaustive search approach. Firstly, a classifier trained from incrementally computable descriptors is used to estimate the class-conditional probabilities of Extremal Regions; Extremal Regions corresponding to local maximum of probabilities in the ER inclusion relation are selected. Secondarily, Extremal Regions passed as the first stage is classified as characters and non-characters using more complex features.

The approaches discussed explores the hierarchical structure of MSERs, but have used different methods for estimating the probabilities of MSERs corresponding to characters. To deal with the big number of repeating MSERs, they have used relevant features in snipping. Next problem with MSER-based approaches, or generally, connected component based approaches and hybrid approaches, is the absence of an effective text entrants construction algorithm. The present methods for text entrant's construction fall into two general approaches: rule-based and clustering-based methods. Neumann et al grouped character entrants by meanings of the text line constrain.The basic notion is that characters in a word can be fitted by one or more top and bottom lines. The text line constrain is quite extravagant, but it's too restrictive for infested text, handwritten text and other language model. Carlos et al. constructed a fully connected graph over character entrants, filtered edges by running a set of tests and used the remaining connected sub graphs as text entrants. Chen et al., pair wised character entrants as clusters with constrains on stroke breadth and height difference, and demoralized a straight line to fit the centroids of clusters. They declared a line as text candidate if it connects three or more character entrants. The clustering-based method presented by Pan et al., clusters character entrants into a tree using the minimum spanning tree algorithm with a learned distance metric; text entrants are constructed by cutting off between-text edges with an energy minimization model.

The discussed rule based methods generally require hand-tuned factors, while the clustering-based method is complicated by incorporating post-processing stage, where one has to specify a rather complicated energy model.

III.TEXT DETECTION

Fig. 1 shows certain images of text signs. It is a challenging problem for an automatic sign detection system to detect text from these images because of affine deformations, highlights, shadows, and secularity. We have to deal with these variations. Table I lists the possible effects on text in an image when lighting, orientation, location, and view angle change, where orientation is the angle between the normal of a sign and optical axis of the camera, and the view angle is view scope.

Fig. 1 Example of an unacceptable low-resolution image

In general, a long focus lens has a narrow view angle and a short focus lens has a wide view angle. An auto white-balanced camera will cause additional changes.

TABLE I EFFECTS ON TEXT IN AN IMAGE CAUSED BY VARIOUS IMAGING CONDITIONS

Text in an image / Imaging condition	Size	Affine	Intensity	Color	Highlight
Lighting	No	No	Yes	Maybe	Yes
Distance	Yes	Yes	No	No	No
Orientation	Yes	Yes	Yes	Yes	Yes
View angle	Yes	Yes	Maybe	Maybe	Yes

By incorporating a number of key improvements than conventional MSER-based methods, we propose a novel MSER-based scene text detection method. The structure of the proposed system and the sample result of each stage is presented in Fig.2.

The proposed scene text detection method contains the following phases:

1) Character candidates extraction. Character candidates are extracted using the MSERs algorithm; maximum numbers of the repeating components are removed by the proposed MSERs pruning algorithm by minimizing regularized variations.

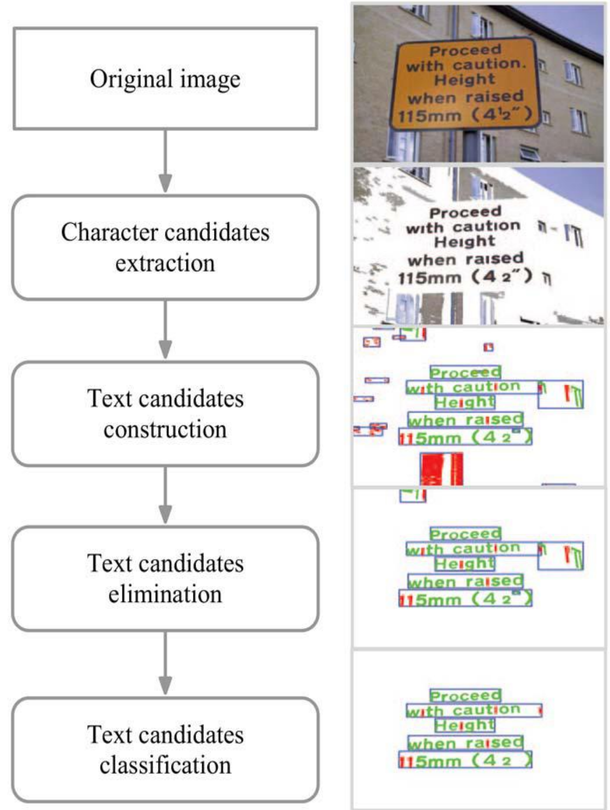

Fig.2. Flowchart of the proposed system and results after each step of the sample. Text candidates are labelled with blue bounding rectangles; character candidates identified as characters are green colored, and others are red colored.

2) Text candidates construction. Distance weights and clustering threshold are learned simultaneously using the proposed metric learning algorithm; character candidates are grouped into text candidates by the use of single-link clustering algorithm using the learned parameters.
3) Text candidates elimination. The posterior probabilities of text candidates correspond to non-texts are assessed using the character classifier and text candidates with high non-text probabilities are removed.
4) Text candidates classification. Text candidates matching to true texts are identified by using the text classifier. An AdaBoost classifier is trained to decide whether an text candidate corresponding to the true text or not.

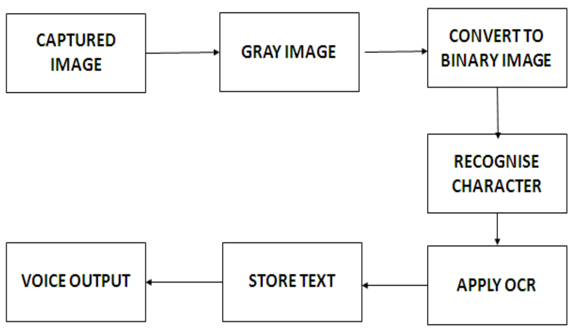

Fig.3. Block Diagram

Consists of a USB camera which captures the product label when brought in front of it. The captured image is an RGB image. This RBG image is converted to gray scale image and finally to binary image which is composed of 0s and 1s. Captured image is processed using MATLAB. Using OCR the characters are recognised from the saved templates. The recognised word is stored in an empty matrix and compared with the database. When the stored word match with the word in the database, there will be similarity and hence the product name will be an audio output in the ear-phones worn by the blind person.

IV.INTENSITY-BASED OCR

A sign can contain graphic or text content. In this research, we focus on the text signs only. OCR for the characters captured from natural scenes faces more challenges than that of document analysis. For a conventional document analysis task, a scanner with a stable embedded lighting system is used to obtain high quality images, which are then easily binarized. For a sign recognition task, however, because the sign image is captured by a camera from natural scenes under various lighting conditions, the signal to noise ratio (SNR) is much lower. If we use binary features for OCR, we cannot promise effectively removing noises before the binarization processing. Although by carefully selecting color spaces we can reduce noises to a certain degree, we have no way to distinguish noises and useful information within the image

Fig. 4. Comparison of (a), (c) binary characters and (b), (d) gray scale

The binarization processing will weigh noises and useful information the same. Furthermore, the segmentation of foreground and background cannot be perfect because of noises. Fig. 8 illustrates some examples from both binary and intensity images, where the color space is carefully selected and hue is used to obtain both binary and intensity images. It is obvious that the binary images are much noisier, while the intensity images have less of a problem separating the background and foreground. Thus intensity images have all the information vital to the decision making stage. Wang and Pavlidis showed the advantages of direct feature extraction from gray scale image for OCR. Our experiments also indicate that the intensity based OCR has advantages than binary OCR for images with low SNR. To avoid irretrievably losing information during the binarization processing, we use the intensity character image directly for feature extraction.

A. Preprocessing:
The characters captured by a camera from natural scenes vary in size, font style, color, and contrast. Furthermore, a character may vary with an affine deformation if the

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

optical axis is not perpendicular to the character plane. We use a scaling algorithm to normalize the size of the character images from different signs. Color variation does not cause a major problem in the OCR phase because we don't use a color image directly. In fact, signs are usually designed with high contrast in both color and gray scale images. We can seldom find a sign using pure colors in both foreground and background, e.g., red characters on a green background.

Lighting sources in natural scenes cause another variation in sign recognition. For example, the sun can cast a highlight point on a sign, and the location of the point will change with time. In addition, many other factors, such as multiple lighting sources and the reflective properties of the surface, will cause uneven intensity distribution of the foreground and background. More specially, the intensity distribution of all strokes within a text region changes in a large dynamic scope: some have obvious contrast, some not, and some are highlighted while others may be dark. We utilize a localized intensity normalization method before feature extraction to reduce intensity distribution changes.

B. Feature Extraction

We used a Gabor wavelet for feature extraction. Because of its superior mathematic properties, Gabor wavelet has been widely used for data compression, face recognition, texture analysis, handwriting recognition and other image processing tasks in recent years. In OCR applications, Gabor wavelet has been applied to binary images, and recently applied to a video stream.

Yoshimura and his colleagues even report using Gabor for feature extraction and linear vector quantization (LVQ) used for feature selection from a video stream. Gabor wavelet is a sinusoidal plane wave with a particular frequency and orientation, modulated by a Gaussian envelope. It can characterize a spatial frequency structure in the image while preserving data about spatial relations, and hence it is suitable for extracting orientation-dependent frequency contents from patterns.

C. Feature Transforms and Recognition

We would like to reduce the number of dimensions of feature vectors because they are computationally expensive and because not all of them are effective for recognition. LVQ and LDA are two common tools for dimension reduction. We use LDA in this research because it can be used not only for dimension reduction, but also for feature optimization.

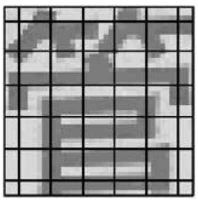

Fig 5. Regions for feature extraction

Fig.6.Examples of text detection from video images

Fig.7. Examples of sign detection.

Fig 8. Example of automatic sign detection and recognition

Fig.8 illustrates the process from detection to recognition. Fig. 8(a) is the original input image, and (b) is the combination of detected candidates from two different resolutions. Fig. (c) is the detection result without affine rectification. It can be seen that only part of the characters are detected in (c).

Similar to the example in Fig. (c), no rectangle sign frame can be found within the image. Therefore, the lines are fit from text. Note that the higher edge intensity corners are used for fitting in Fig. (d). The detected result with affine rectification is in Fig. (e), where all characters are detected and all Chinese characters are recognized correctly.

V.SIMULATION RESULTS

Fig. 9. Input Image

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

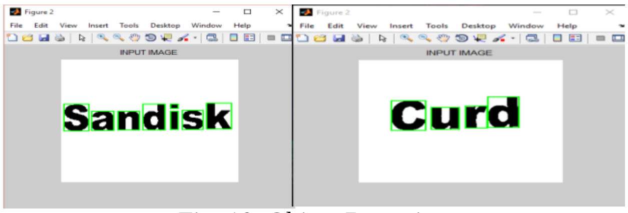

Fig. 10. Object Detection

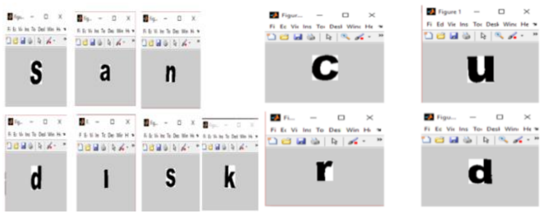

Fig. 11. Object Extraction

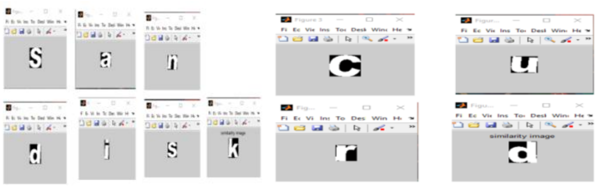

Fig. 12. Character after Recognition

The extracted characters of SANDISK is matched with the recognised characters, then the final output is a voice output as “SANDISK”. The extracted characters of CURD are matched with the recognised characters, and then the final output is a voice output as “CURD”. Characters

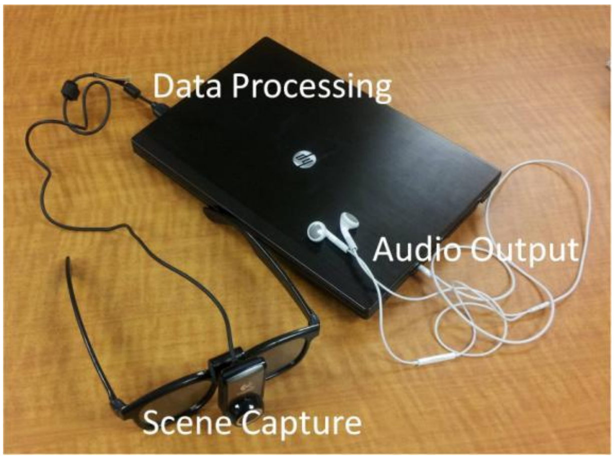

Fig.13. Prototype of the Proposed System

VI.CONCLUSION

This paper presents a new MSER-based scene text detection method with several novel techniques. we proposed a fast and accurate MSERs pruning algorithm that enables us to detect most characters even when the image is in low quality and then a novel self-training distance metric learning algorithm that can learn distance weights and clustering threshold simultaneously; text candidates are constructed by grouping character candidates by the single-link algorithm using the learned parameters. Then, we put forward to use a character classifier to estimate the posterior probability of text candidate related to non-text and remove text candidates with high non text probability, which helps to build a more powerful text classifier. Finally, by integrating the above new techniques, we build a accurate scene text detection system that exhibits superior performance over state-of-the-art methods on a variety of public databases. We empirically analyze several main limitations of our technology for further research, how to detect highly blurred texts in low resolution natural scene images is a near future issue. Second, some multilingual texts have quite various characteristics from English texts. Therefore, adaptively detecting a variety of multilingual texts simultaneously is also a challenge. Third, for similar multiple text lines with a seriously skewed distortion. Internet: By harnessing the capability to access Internet through the ARM-based computer, we can provide a wider range of applications and access to information than go beyond the current in-person communication. Applications: Along with Internet access, the users could install custom applications that can be developed by anyone and made available online. Examples could range from new languages (for speech synthesis, braille contractions, etc.)to SMS messaging, even books or games.

REFERENCES

[1] Y. Zhong, H. Zhang, and A. Jain, “Automatic caption localization in compressed video,” IEEE Trans. Pattern Anal. Mach. Intell.,vol. 22, no. 4, pp. 385–392, Apr. 2000.

[2] H. Li, D. Doermann, and O. Kia, “Automatic text detection and tracking in digital video,” IEEE Trans. Image Process., vol. 9, no. 1, pp. 147–156, Jan. 2000. P.Devi, B.Saranya, L.Malathi, “Fingerprint combination by minutiae and coordination extraction for privacy protection” in International Journal of Latest Trends in Engineering and Technology, vol.6, issue3, pp. 595-601,2016.

[3] J. Weinman, E. Learned-Miller, and A. Hanson, “Scene text recognition using similarity and a lexicon with sparse belief propagation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 10,pp. 1733–1746, Oct. 2009.

[4] X.-C. Yin, H.-W. Hao, J. Sun, and S. Naoi, “Robust vanishing point detection for mobile cam-based documents,” in Proc. ICDAR, Beijing, China, 2011, pp. 136–140.

[5] K. Kim, K. Jung, and J. Kim, “Texture-base approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp. 1631–1639, Dec. 2003.

[6] B. Epshtein, E. Ofek, and Y. Wexler, “Detecting text in natural scenes with stroke width transform,” in Proc. IEEE Conf. CVPR, San Francisco, CA, USA, 2010, pp. 2963–2970.

[7] J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide baseline stereo from maximally stable extremal regions,” in Proc. Brit.Mach. Vis. Conf., vol. 1. 2002, pp. 384–393.

[8] A. Shahab, F. Shafait, and A. Dengel, “ICDAR robust reading competition challenge 2: Reading text in scene images,” in Proc. ICDAR, 2011, pp. 1491–1496.

[9] H. Chen et al., “Robust text detection in natural images with edgeenhanced maximally stable extremal regions,” in Proc. IEEE Int. Conf. Image Process., 2011, pp. 2609–2612.

[10] C. Merino-Gracia, K. Lenc, and M. Mirmehdi, “A head-mounted device for recognizing text in natural scenes,” in Proc. Int. Workshop CBDAR, Beijing, China, 2011, pp. 29–41.

[11] L. Neumann and J. Matas, “Real-time scene text localization and recognition,” in Proc. IEEE Conf. CVPR, Providence, RI, USA, 2012,pp. 3538–3545.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 3, Issue 11, November 2016

[12] L. Neumann and J. Matas, “Text localization in real-world imagesusing efficiently pruned exhaustive search,” in Proc. ICDAR, Beijing, China, 2011, pp. 687–691.
[13] L. Neumann and J. Matas, “A method for text localization and recognition in real-world images,” in Proc. ACCV, vol. 3.Queenstown, New Zealand, 2011, pp. 770–783.
[14] C. Shi, C. Wang, B. Xiao, Y. Zhang, and S. Gao, “Scene text detection using graph model built upon maximally stable extremal regions,” Pattern Recognit. Lett., vol. 34, no. 2, pp. 107–116, 2013.
[15] K. Jung, K. Kim, and A. Jain, “Text information extraction in images and video: A survey,” Pattern Recognit., vol. 37, no. 5, pp. 977–997, 2004.
[16] J. Liang, D. Doermann, and H. Li, “Camera-based analysis of textand documents: A survey,” IJDAR, vol. 7, no. 2–3, pp. 84–104,2005.
[17] J. Zhang and R. Kasturi, “Extraction of text objects in video documents: Recent progress,” in Proc. IAPR Workshop DAS, Nara,Japan, 2008, pp. 1–13.
[18] X. Yin, X.-C. Yin, H.-W. Hao, and K. Iqbal, “Effective text localization in natural scene images with MSER, geometry-based grouping and ada boost,” in Proc. Int. Conf. Pattern Recognit., Tsukuba, Japan, 2012, pp. 725–728.
[19] T. Q. Phan, P. Shivakumara, and C. L. Tan, “Detecting text in the real world,” in Proc. ACM Int. Conf. MM, New York, NY, USA, 2012, pp. 765–768.
[20] K. Wang, B. Babenko, and S. Belongie, “End-to-end scene text recognition,” in Proc. IEEE Int. Conf. Comput. Vis., Barcelona, Spain, 2011, pp. 1457–1464.

BIOGRAPHIES

Ms. P. Devi received her M.E degree in Communication Systems from Sri Krishna College of Engineering and Technology, Anna University, Coimbatore in the year 2011. She received her B.E. degree in Electronics and Communication Engineering from VLB Janakiammal College of Engineering and Technology, Anna University, Chennai in the year 2009. Currently she is working as Assistant Professor in ECE Department at Sri Ramakrishna Institute of Technology .She is having five years of teaching experience. She has presented papers in National Conferences. She has published her papers in National and International Journals. Her fields of interest are Digital Image Processing and Wireless Networks.

Ms. B. Saranya completed her M.E. degree in Embedded Systems at Kumaraguru College of Technology, Anna University, Chennai in the year 2015. She received her B.E. degree in Electronics and Communication Engineering from Sri Ramakrishna Institute of Technology, Coimbatore, Anna University, Chennai in the year 2013. Currently she is working as Assistant Professor in ECE Department at Sri Ramakrishna Institute of Technology. She is having 17 months of teaching experience. Her area of interest is Embedded System Design. She has published papers in International Journals.

Ms. B. Abinayaa completed her M.E. degree in Embedded Systems at Sri Ramakrishna Engineering College, Anna University, Chennai in the year 2016. She received her B.E. degree in Electronics and Communication Engineering from Sri Shakthi Institute of Engineering and Technology, Coimbatore, Anna University, Chennai in the year 2013. Currently she is working as Assistant Professor in ECE Department at Sri Ramakrishna Institute of Technology. She is having 6 months of teaching experience. Her area of interest is Embedded System Design. She has published papers in International Journals.

Ms. G. Kiruthikamani completed her M.E. degree in Embedded System Technologies at Sri Ramakrishna Engineering College, Anna University, Chennai in the year 2015. She received her B.E. degree in Electronics and Communication Engineering from Sri Ramakrishna Engineering College, Coimbatore, Anna University, Chennai in the year 2013. Currently she is working as Assistant Professor in ECE Department at Sri Ramakrishna Institute of Technology. She is having 6 months of teaching experience. Her area of interest is Embedded System Design and Microcontroller. She has published papers in International Journals.

Ms. N. Geethapriya received the B.E degree in Electronics and Communication Engineering from Avinashilingam University, Coimbatore, India, in 2015. Currently, she is a Post Graduate Student with the Department of Electronics and Communication Engineering, Sri Ramakrishna Institute of Technology, Coimbatore. Her current research interests include Wireless communication, Internet of Things.