Criterion 3- Research, Innovations and Extension
3.1- Resource Mobilization for Research

3.1.1 - Grants received from Government and non-governmental agencies for research projects / endowments in the institution during the last five years

Sr. no.	Name of the Project/ Endowments, Chairs	Name of the Principal Investigator/Co-investivator	Year of Award	Amount Sanctioned	Page no
1	System for Secure Fine-Grained Access Control and Authorization of Digital Assets and Operations	PI: Dr. Sunil Dhore Co-PI: Dr. Sagar Rane	2021-23	745.78 Lakhs	**3**
2	Text Detection and Recognition from Indian Street Scene Images	Dr. R Jayadevan	2016-17	14.50 Lakhs	**12**

3.1.1.1 : Total Grants from Government and non-governmental agencies for research projects / endowments in the institution during the last five years (INR in Lakhs)

Year	2022-23	2021-22	2020-21	2019-20	2018-19
Number	**745.78**	**749.28**	**10**	**0**	**14.5**

Principal
Army Institute of Technology
Dighi Hills, Pune - 411015

Brig Abhay A Bhat (Retd)
Director
Army Institute of Technology
Dighi Hills, Pune - 411 015

Work in Progress record
AY 2022-23

Sr No	Name of the Project/ Endowments, Chairs	Name of the Principal Investigator/ Co-investigator	Department of Principal Investigator	Year of Award	Amount Sanctioned	Duration of the Project	Name of the Funding Agency
1	System for Secure Fine-Grained Access Control and Authorization of Digital Assets and Operations	PI: Dr. Sunil Dhore Co-PI: Dr. Sagar Rane	Computer Engineering	2021-22	Rs. 7,45,78,000/-	2 Years	National Security Council Secretariat

Principal
Army Institute of Technology
Dighi Hills, Pune - 411015

Brig Abhay A Bhat (Retd)
Director
Army Institute of Technology
Dighi Hills, Pune - 411 015

RESTRICTED

National Security Council Secretariat

2nd Floor, Sardar Patel Bhawan,
Sansad Marg, Parliament Street
New Delhi – 110001

Subject: **Meeting of the 1st Project Review & Monitoring Committee (PRMC) of project titled "R&D of System for Secure Fine-grained Access-Control & Authorisation of Digital assets and operations" held on 20th May, 2022**

Please find enclosed minutes of the 1st Project Review and Monitoring Committee (PRMC) meeting of the project titled "R&D of System for Secure Fine-grained Access-Control & Authorisation of digital assets and operations" which was held on 20th May, 2022. The meeting was conducted in hybrid mode under the Chairmanship of Maj Gen Manjeet Singh, Joint Secretary, NSCS.

(R. C. Sharma)
Director
Tel: 011-23451576
Email id: rc.sharma70@gov.in

Enclosure: (07 pages)

1. **Dr. Rajesh Pillai**, Scientist, Director (SAG), Scientific Analysis Group (SAG), DRDO, Timarpur, Metcalf House Complex, Civil Lines, Delhi – 110054
2. **Dr. Somnath Chandra,** Scientist F, Cyber Security R&D Division, MeitY Electronics Niketan, 6 CGO Complex, New Delhi - 110003
3. **Dr. Pradeepthi VK**, Professor & Researched, CR Rao Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS), University of Hyderabad Campus, Gachibowli, Hyderabad – 500046.
4. **Prof. NV Narendra**, Professor & Researched, Institute of Development and Research in Banking Technology (IDRBT), Hyderabad - 500057
5. **Shri Jivan Kumar**, ADD, NCIIPC, NTRO, Block-III, Old JNU Campus, New Delhi-110067
6. **Wg. Cdr JM Cladius**, DCyA, Room No 590, D Ops (IEW), Air HQ (VB), Rafi Marg, New Delhi - 110011
7. **Shri Deepak Gupta**, Assistant Director, IB, 35 SP Marg, New Delhi -110021
8. **Shri SK Jha**, DCIO, IB, 35 SP Marg, New Delhi -110021

NSCS ID No. 46/167/2021-NSCS(CS) Dated: 30 Aug, 2022

Copy to:

1. **NCSC, NSCS**
2. **AS(RA), NSCS**
3. **JS(NN), NSCS**
4. **JS(MS), NSCS**
5. **Prof (Dr) Sunil Dhore**, Professor & Head of the Department, Computer Engineering, Army Institute of Technology (AIT), Dighi Hills, Pune – 411015, Maharashtra
6. **Shri Nilesh Dhande**, 42 Labs, 2nd Floor, C-1-A, Baner, Pune – 411045, Maharashtra.

NATIONAL SECURITY COUNCIL SECRETARIAT

Subject: Minutes of the 1st Project Review & Monitoring Committee (PRMC) meeting of the Project "R&D of System for Secure Fine-grained Access-Control & Authorisation of digital assets and operations".

JS(MS) Chaired the meeting of the "Secure Fine-grained Access-Control & Authorisation" project on 20 May 2022. The project is being jointly implemented by Army Institute of Technology (AIT) and M/s 42 Labs. The meeting was conducted in hybrid mode. List of the participants is annexed as **Annexure-I**.

2. The Chairman welcomed the members and told about importance of the project. He informed the PRMC members that the project was approved at total outlay of Rs 7.4578 crore for a period of two years and an initial amount of Rs 2.37 crore has been released to the executing agency. He then asked Project Investigator (PI) Prof Sunil Ghore, of Army Institute of Technology to give presentation on the progress of the project as per the time-lines of the approved project.

3. PI initially gave brief of the project and conveyed that digital assets are growing at an exponential rate. Most of the physical assets have digital representation and, in some case have controls too (eg., Industrial IoT, Connected Cars, Drones etc.). The traditional enterprises' digital boundaries are also rapidly fading as situations demand that users need to either access or control digital assets from anywhere. The traditional access control mechanism where one needs to solely depend on "trusted" middleware to store and grant access is now insufficient to tackle the challenge.

4. In every interaction related to access and authorization; security, trust and hence provability are the key requirements irrespective of type of asset or type of access. Thus, cryptographically enabled and enforced access control becomes highly relevant. The proposed system should satisfy following requirements: -

a. Trust: automation of the access-control process using advanced information-theoretic cryptography, thus removing the need for trust on a middleman (admin, software) for ensuring correct grant of access to assets.

b. Fine-grained: capability of asset owner(s) to specify a variety of restrictions that narrows down the conditions under which access is permitted, in terms of – asset, subject and context.

c. Access-Control: restriction of access to objects worth protecting (assets – could be data/operations).

d. Identity-based: a subject gaining access to an asset, should do so while complying with access-control restrictions associated with all concerned identities – asset owner(s), the subject, and the restriction policy

5. The PI conveyed that the project aims to build the protocol and inter-disciplinary platform which will provide highly secure and provable access-control and authorization for consumption by variety of systems.

6. Objective of the project is to design algorithm/protocol and software platform to-

a. Grant context sensitive and provable access to a set of actors (apps/APIs/Humans/devices) and allow flexibility in specifying the access rights.

b. Construct composite access policies consisting of multiple actors, attributes and constraints

c. Provide facility to include, apart from static attributes, Dynamic constraints (e.g., location/IP/authorised device compliance etc.), Dynamic policy modification without impacting the secured asset

d. Preserve privacy and integrity of a policy, and

e. Support multi-authority model where constrains from different providers must be admissible.

7. The project has been planned in three phases of eight months each project conception and initiation, system development, and testing and production. Pert Chart of implementation of all three phases is given at **Annexure-II**.

8. Phase-I of the project i.e. project conception and initiation which includes requirement analysis, detailed solution design, rule engine conceptualization, capital equipment purchase proceedings, and policy instrumentation interface have been completed. Details of the goals of Phase-I and status of deliverables may be seen at **Annexure-III.**

9. PI conveyed that Phase-II of the project involves policy instrumentation interface, identity layer and policy storage, policy crypto library, policy enforcer module, policy validation/evaluation, policy evidence store interface, data source layer development, and functional testing. He conveyed that the work of Phase-II is yet to be started.

10. PI explained expenditure of the initial release of Rs 236.885 lakh and requested release of Rs 286.09 lakh for Phase-II of the project. **The PRMC recommended release of Rs 286.09 lakh subject to submission of Utilisation Certificate of the fund released in Phase-I.** The executing agency was also told to submit the interest amount of the already released fund for the project.

11. **PRMC Observations and Comments:**

a. The PRMC observed that the project is heading in the right direction.

b. PRMC commented that data for which access needs to be controlled cannot always be centralised. PI conveyed that there is a provision for distributed data store in the design and development efforts can be undertaken on use case basis.

c. On catering to the negative policies which contradicts, PI informed that policy system design has considered prohibitions (explicit denying of operations or access based on conditions), and further research and enforcing them is work in progress.

[Action by AIT & 42 labs]

d. PRMC advised that negative policies would be tricky, therefore, the team needs to think through them while its implementation; PI conveyed that approach to access is whitelisting based, so need for negative policies should not arise.

e. On query regarding the conflict resolution policies that are utilized, PI conveyed that the project team will explore this issue.

[Action by AIT & 42 labs]

f. PI apprised the PRMC that the solution will be implemented at NSDL (User Agency). PRMC discussed that NSDL has huge database on Prim, if they can move a fraction of data into the cloud then they will be able to save a lot of costs and the current solution could enforce access control on cloud infra without compromising security. The PRMC discussed about data archival & key management. PI informed that requirement analysis is in progress.

[Action by AIT & 42 labs]

g. Executing Agency will discuss the Personnel Identifiable Information (PII) use case of NSDL and explain this to PRMC members separately.

[Action by AIT & 42 labs]

h. PRMC suggested AIT-FT 42 Labs to build a list of uses cases in multiple industries. PI conveyed that the matter will be discussed with various industry experts and come up with the comprehensive list.

[Action by AIT & 42 labs]

i. AIT-FT 42 Labs should plan for the roadmap post-implementation at the user agency for the sustenance of the solution and project benefits.

[Action by AIT & 42 labs]

j. AIT-FT42 Labs should have a working solution deployed into the user agency to make this a success.

[Action by AIT & 42 labs]

k. PRMC recommended collaborating with IDRBT and NCIIPC members for use case deliberation.

[Action by AIT & 42 labs]

I. PRMC advised that benchmarking and evaluation of the solution be taken from an independent partner or from SAG.

[Action by AIT & 42 labs]

12. The meeting ended with the vote of thanks to the Chair.

****

NATIONAL SECURITY COUNCIL SECRETARIAT

ANNEXURE-I

LIST OF ATTENDEES

Sr.No.	Name of the officer/Designation	Organisation
1.	Maj Gen Manjeet Singh, Joint Secretary - Chairman	NSCS
2.	Shri G Narendra Nath, JS	NSCS
3.	Dr Rajesh Pillai, Scientist, Director, SAG	DRDO
4.	Sqn Ldr Ravin, IDS HQ	DCyA
4.	Shri R C Sharma, Director	NSCS
5.	Shri Jivan	NCIIPC
6.	Shri Deepak Gupta, AD	IB
7.	Shri SK Jha, DCIO	IB
8.	Prof NV Narendra, Prof & Researcher	IDRBT
9.	Smt Pradeepthi KV, Prof & Researched	CR Rao Institute

ID	Work Item \ Phase & Month	PHASE ONE							PHASE TWO								PHASE THREE							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	**Project Conception and Initiation**																								
1.1	Requirements Analysis																								
1.2	Detailed solution design																								
1.3	Rule Engine conceptualization																								
1.4	Purchase Proceedings																								
1.5	Policy Instrumentation UI																								
2	**System Development**																								
2.1	Identity Layer and Policy Storage																								
2.2	Policy Crypto Library																								
2.3	Policy Enforcer Module																								
2.4	Policy Validation/ Evaluation																								
2.5	Policy Evidence Store Interface																								
2.6	Data Source Layer Development																								
2.7	Functional Testing																								
3	**Various Testing, Production**																								
3.1	Distributed Blockchain setup																								
3.2	IoT Use Case																								
3.3	Performance Testing																								
3.4	Security Analysis & Testing																								
3.5	iterative dev-test sprints																								
3.6	Production Deployment																								
3.7	Users training																								

Annexure-III

Area	Activities	Status
Team Building	Training Presentations	Completed
	Research Papers repository	Completed
	Onboard Student interns	Completed
Cryptosystem Conceptualization	Policy Cryptosystem Draft Design	Completed
	Problem Statement Solution Documents for ▪ Dynamic Access Policy Revocation ▪ Generalization of Access Policy ▪ Policy Security ▪ Fuzzy Attributes ▪ Multi Authority	Completed Completed Completed Completed Completed
	TCG Crest solution document review and security analysis report	Completed
Rule Engine Conceptualization	XACML Demo Presentations	Completed
	NGAC Demo Presentations	Completed
Purchase Proceedings	RFP Process with Equipment Specification	Completed
	Hardware Procurement & Installation	Completed
Use Case Req. Analysis	High-Level Use Case Requirements Analysis draft for Key management	Completed

Grants Received from Government Agencies

AY 2019-20

Sr No	Name of the Project/ Endowments, Chairs	Name of the Principal Investigator/ Co-investigator	Department of Principal Investigator	Year of Award	Amount Sanctioned	Duration of the Project	Name of the Funding Agency
1	Text Detection and Recognition from IndianStreet Scene Images	Dr. R Jayadevan	Information Technology	2016-17	Rs 14,50,000/-	3 Years	DST Science and Engineering Board

Principal
Army Institute of Technology
Dighi Hills, Pune - 411015

Brig Abhay A Bhat (Retd)
Director
Army Institute of Technology
Dighi Hills, Pune - 411 015

2. SERB

PROGRESS REPORT

1. Project Title: Text Detection and Recognition from Indian Street Scene Images	DST No: YSS/2015/000812
2. PI (Name & Address): Dr. Jayadevan R, Associate Professor, Army Institute of Technology, Dighi Hills, Pune, Maharashtra, Pin:411015	Date of Birth: 18-02-1981
3. Co-PI (Name & Address): Not Applicable	Date of Birth: Not Applicable

4. Broad area of Research
Image processing and Pattern recognition.

4.1 Sub Areas
Mobile camera based image processing; Scene text detection; Scene text localization; Scene text extraction; Scene text recognition; Dataset development; Devanagari scene text processing; Multilingual scene text processing

5. Approved Objectives of the Proposal :

In a multilingual country like India, the local language (script) is used to write/print signboards, bus-route details, instructions, banners, wall-writings, road-names, shop-names, apartment-names, advertisements etc. This creates difficulties for tourists in getting useful information about route, direction, names, instructions etc. A person residing in a place may only be proficient in his/her native language (script), whereas a foreigner visiting India may noteven know any Indian script. It would be useful if she/he could use a portable imaging deviceto snap the text and some software would recognize and translate it into a language known to him/her. The proposed project aims at doing the first part, i.e. camera based recognition of Devanagari text from street scene images.

Devanagari script is used to write Hindi or Marathi texts. Such texts are abundant in outdoor urban scenes in Hindi and Marathi speaking belts of India. Here the images taken by digital cameras (attached to cellular phones) are only concerned. Such phones are highly mobile, easy to use and even a cheap version has a good camera resolution. It is possible to do limited computation in reasonable time within the device or getting it done in a remote server. Thus, a cellular phone with camera based system shall be very useful for people from different states and abroad who cannot read the script. Some major challenges of such a system are low resolution, uneven lighting, perspective distortion, non-planar surfaces, complex back grounds, climatic conditions, complex foreground and multi-script foreground.

The following are the main objectives of the project:

a) The project aims at developing a system capable of detecting and recognizing Devanagari text from Indian street scene images.
b) Such a system capable of translating Devanagari text from outdoor scene images will be useful in sign detection and translation for tourists, understanding environment for blind persons, automatic navigation systems, content-based image search, reading license plates, name plates, banners and instructions, object recognition and scene understanding.

c) There is no data-set available for experimentation related to scene text detection, extraction and enhancement in Indian context. A bench-marking data-set containing at least 5000 street scene images has be created and the same has to be shared with global research community.

Date of Start: 15th February 2016.	Total amount received: Rs. 10,53,333
Date of completion: 14th February 2019	Expenditure as on 31-03-2019:Rs. 5,81,724

6. Methodology:

The methodology adopted for the work is shown in the figure below. The first step to be completed is a comprehensive literature survey, which will identify some potential techniques for Devanagari text detection, extraction, enhancement and recognition. The identification of potential techniques shall be based on three parameters namely precision, recall and F-measure.

The characteristics of Devanagari script will also play a significant role in selecting a technique(s). It is to be noted here that some techniques which are highly successful in detecting Latin texts from scene images may not do well in detecting Devanagari text.

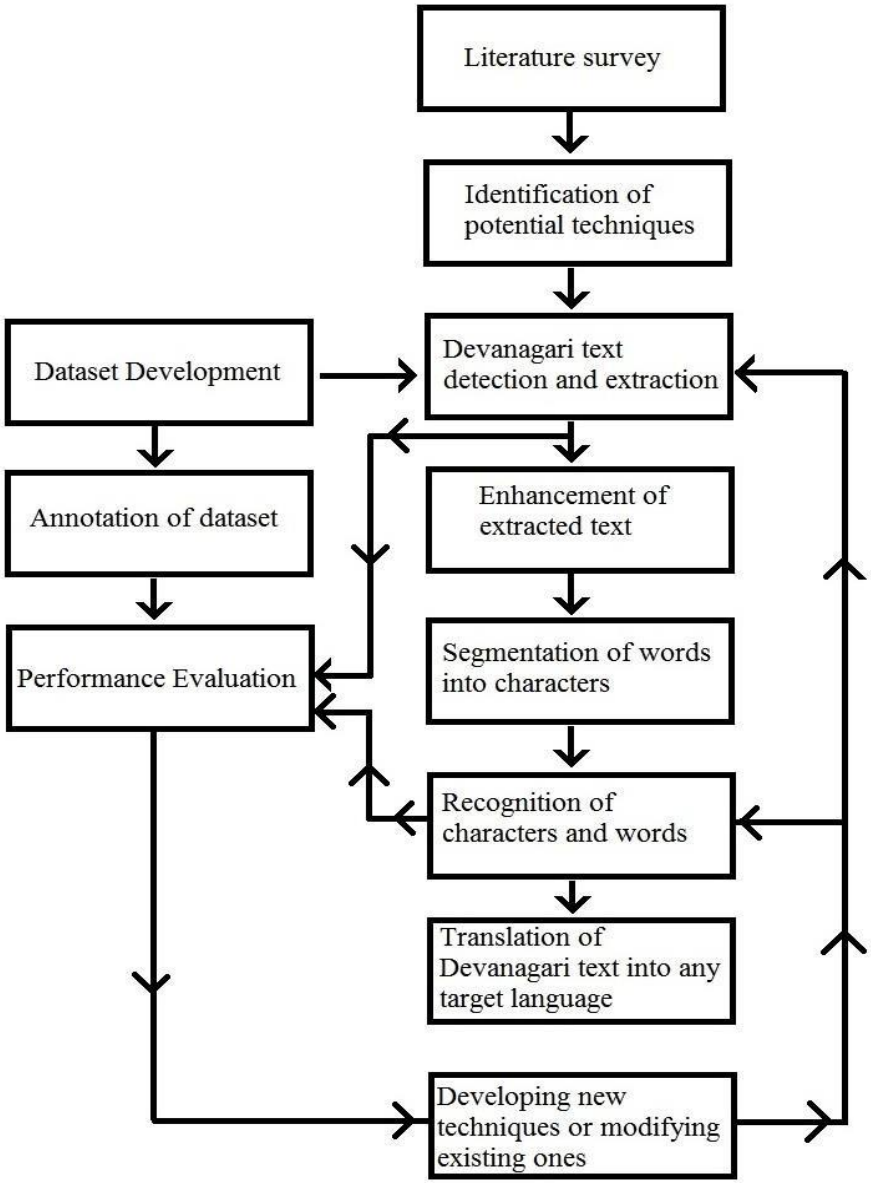
A benchmarking dataset is to be created for experimentation and the same has to be shared with the research community in India/Abroad. Such a dataset should have at least 5000 images in it with various levels of complexity and resolution.

Dataset annotation is the process of creating ground truth transcriptions in the form of tag files. Such tag-files shall contain the location information of the text (words) in the images as a set ofrectangles. It can also contain the segmentation points along with the ground truth constituent characters of each word. These tag files will be useful in automatic evaluation of the performance of text detection and text recognition algorithms.

Text extraction is the stage where the text components are to be segmented from the background. Enhancement of the extracted text components is required as the text region usually is of low-resolution and is prone to noise. Also, skew and slant corrections can be done at this stage to improve the recognition accuracy.

The extracted text (word) image has to be converted to a binary image and enhanced before it isfed into a recognition engine. Each and every word or string has to be segmented into its constituent characters or symbols for classification and recognition. Character recognition will further lead to the recognition of words and text as a whole. The entire text has to be successfully recognized to make it ready for translation into other languages. Existing language translation tools like Google-Translate can be utilized for this purpose.

The precision and recall are used to measure the retrieval power. Precision is defined as the number of correct estimates divided by the total number of estimates. Recall is defined as the number of correct estimates divided by the total number of targets. The F-measure combines the precision and recall figures into a single measure.



The performance evaluation of the text detection and recognition algorithms are important to ensure higher precision, recall, F-measure and recognition accuracy. Certain targets shall be set for text detection and recognition. The algorithms shall be modified or new ones shall be developed if they fail to achieve the targets.

7. Salient Research Achievements:

7.1 Summary of Progress

a. A benchmarking dataset containing more than 5000 street scene images has been created. Images containing banners, boards, wall paintings, bus route details, nameplates written in Devanagari script were captured using a mobile camera. The ground truth information of 1250 images is created for automatic evaluation of the proposed techniques. This benchmarking dataset will be made available to theglobal research community soon.

b. To detect Devanagari text from scene images, two techniques are proposed. The first technique is based on edge detection and the second is based on colour information. Later the two techniques have been combined to achieve higher accuracy. Edges play an important role in multi sized, multi colored and multi oriented text identification. In the edge-based technique, image is preprocessedbefore applying the edge detector. Region-based properties are then used for the elimination of irrelevant edges. Text regions are then identified by connected component analysis. Colour homogeneity is also an important feature for text detection in high contrast images. In the colour-based technique, background and foreground regions are separated by colour based clustering followed by region- based property filtering and connected component analysis.

c. Another technique to detect and localize text written in Devanagari script from scene images is also proposed. Initially, candidate regions are localized using low level features like edge and colour. Due to the complex nature of scene images, these regions may contain irrelevant information. Stroke width transform (SWT)and geometric features are then extracted from these localized regions for correctly identifying the text regions. An efficient technique is proposed for the extraction of stroke width from dark text (foreground) on light background as well as from light text (foreground) on dark background. Methods based on heuristic rules are inefficient for text and non-text identification due to the nonlinearity of extracted features. As a result, Naïve Bayesian Classifier (NBC), Support Vector Machine (SVM) and a Decision Tree Classifier (DTC) are used for classifying text and non-text regions using stroke width transform and geometric features.

7.2 New Observations:

In the edge-based technique, the image is smoothed before applying 'canny' edge detector. There are two main advantages of image smoothing; it reduces enormous generation of edges in case of complex background due to trees, leaves, etc. and it minimizes shadow effect. Image smoothing can be achieved by applying either convolution or filtering. Experimentation is carried out using both and the results obtained are mentioned in Table.1. It is clear from the table that convolution-based smoothing produces better results than Gaussian filter-based smoothing. Overall result states that edge-based technique has good recall but produces large number offalse alarms. Edges are invariant with different light conditions, font styles, multi colored and multi sized fonts. Hence, edge-based technique detects maximum number of text-regions. On the other side, it also produces enormous and irrelevant edges due to complex scene background, shadow effect and poor resolution.

Table 1. Accuracy of edge-based technique

Smoothing Techniques	Precision	Recall	F-measure
Convolution	59.40 %	83.50 %	69.40 %
Gaussian Filter	54.60 %	84.30 %	66.30 %

In the colour based technique, the input image is segmented into text and background clusters by applying K-means in YCbCr colour space whereas L*a*b ismore popular choice of many of the researchers. Hence, experimentation is carried out in both the colour spaces and the evaluated results are mentioned in Table 2.

The advantage of colour-based technique is that there is very less difference between precision and recall. Colour-based segmentation produces clean image (only two colour image) hence false bounding boxes are hardly generated. Improved results are shown in Figure 1. Colour-based technique has good precision compared to edge-based technique.

Table 2. Accuracy of colour-based technique

Colour-based Techniques	Precision	Recall	F-measure
YCbCr Colour Space	67.60 %	69.40 %	68.50 %
L*a*b Colour Space	65.90 %	68.70 %	67.20 %

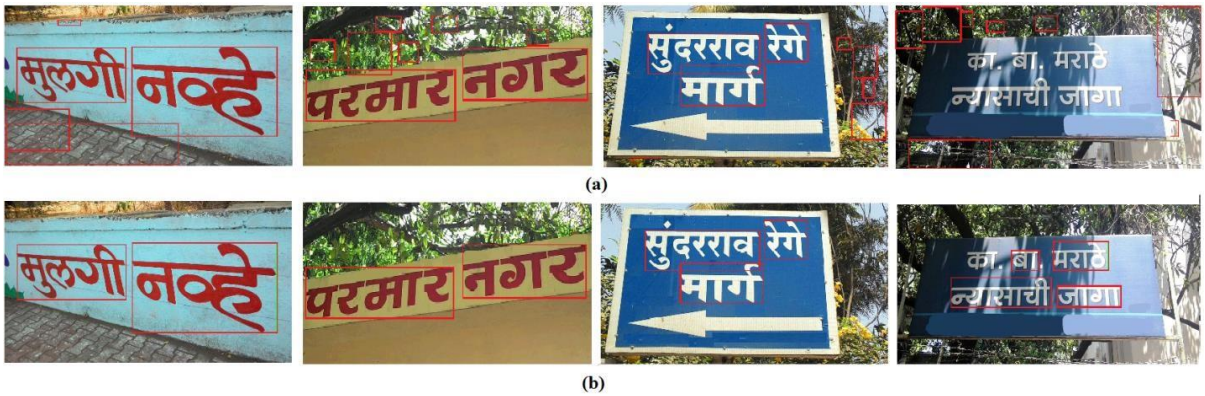


Figure 1. Compararision of edge-based and colour-based techniques on same images. (a) Edge-based technqiue (b) Colour-based technique

Edge-based technique has high recall but detects too many extra annotations whereas colour-based technique generates nearly exact annotations. So, both the techniques are combined to take advantages and overcome limitations of each other. After performing colour-based segmentation, edge map is derived by applying edge detection. This is followed by region-based property filtering and connected component analysis. Smoothing is a major step in edge-based technique but as colour-based technique divides the image into two colours, image appears cleaner and clearer. Therefore, smoothing is not required before applying edge detector. Results with and without applying convolutions are given in Table 3.

Table 3. Accuracy of combined approach.

Combined Approach	Precision	Recall	F-measure
with Convolution	73.12 %	73.93 %	73.52 %
without Convolution	74.05 %	74.46 %	73.75 %
without Region Fill	74.05 %	73.46 %	73.75 %
with Region Fill	74.43 %	74.22 %	**74.32 %**

Similarly, region filling is one of the major steps in CCA in edge-based technique to remove irrelevant regions. Colour-based technique produces solid text cluster hence region filling is not required. While combining both the techniques, results are obtained with and without region fill operation. It is experimentally proved that region fill operation improves efficiency of the technique. Figure 2 shows comparative output of colour-based technique and colour-edge-based technique. In

colour-based technique, cluster is selected based on the assumption that background colour occupies more area than font. But in Fig. 2, font colour occupies more area than background hence wrong cluster is selected for further processing. Hence text boxes are not detected in output image. Whereas in combined approach, aftersegmentation instead of selecting a particular cluster according to the size, 'canny' edge detector is applied directly.

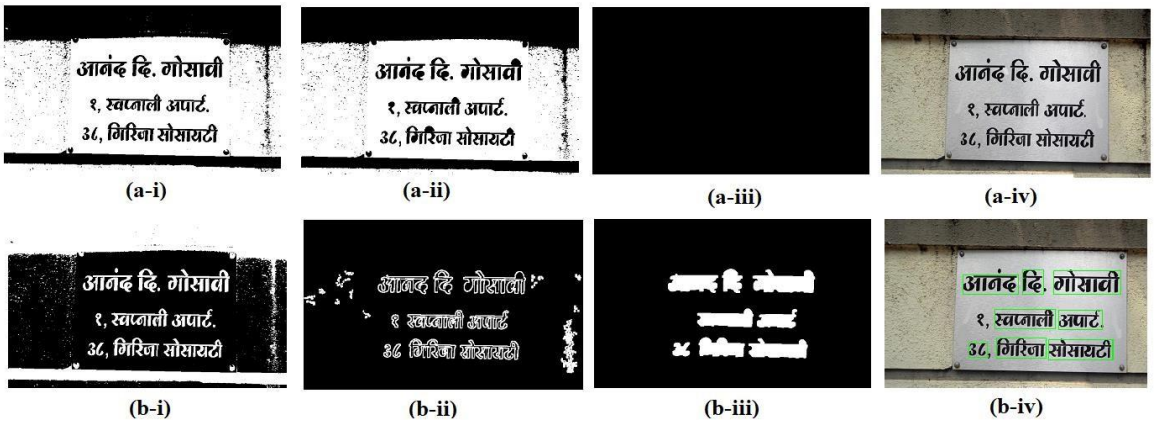


Figure 2. Comparative output of colour-based and colour-edge-based (a) Colour-based technique (b) Colour-edgebased technique ; Steps (i) Selected cluster (ii) Region-based filtering (iii) CC analysis (iv) Detected text

Many scene images of navigation boards have arrows on it to show the direction. Generally, arrows are of the same colour and considered as a part of the information given on a board. Hence it is difficult to differentiate arrows from text. A special step for arrow detection and removal is carried out to improve theaccuracy as shown in Figure 3. Table 4 shows that the performance of the techniqueimproves due to the arrow removal step.

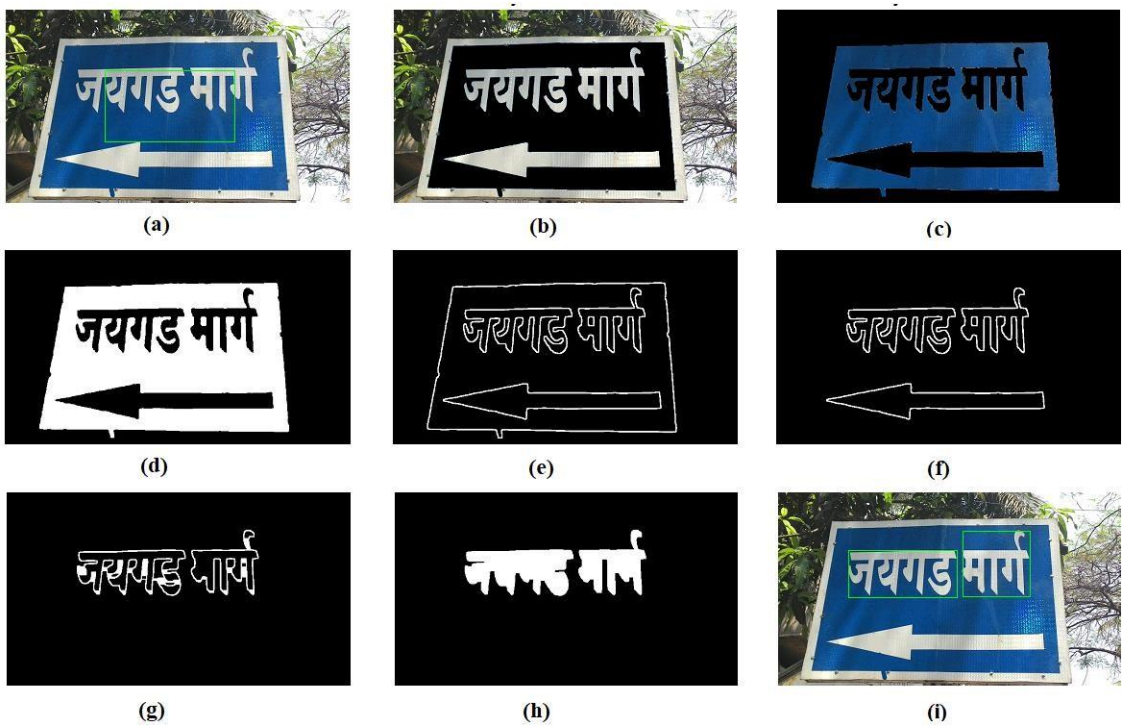


Figure 3. Combined approach with arrow removal (a) Selected center area (b) Objects in cluster one (c) Objects in cluster two (d) Randomly selected cluster (e) Edge detection and gradient mask (f) Region-based filtering (g) Arrow removal (h) Region filling and CC analysis (i) Detected text.

Table 4. Accuracy of combined approach after arrow removal

Combined Approach	Precision	Recall	F-measure
arrow removal	74.43 %	74.22 %	74.32 %
without arrow removal	67.60 %	78.20 %	72.50 %

To the best of my knowledge, there is no any benchmarking dataset available in Devanagari script for experimentation. It is found in literature survey that Bangla and Devanagari words have common characteristics like header-lines. Hence, the results of various existing methods for Bangla and Devanagari are considered and compared in Table 5. The proposed technique is more robust as it is evaluated using a dataset containing 1250 images, whereas others have used comparatively smaller datasets. Though these techniques have not been evaluated using a common benchmarking dataset, by considering the diversity and the number of images in the dataset it is evident the proposed technique is more accurate than the existing ones.

Table 5. Performance of the proposed and existing techniques

Techniques	Script	Features	No. of Images	F-measure
Proposed Technique	Devanagari	Colour, Edge & morphological operations	**1250**	**74.32%**
Raj et al. [30]	Devanagari	Binarization & morphological operations	100	73.49%
Chowdhury et al. [14]	Devanagri & Bangla	Edge, stroke & geometric features	100	72.98%
Sounak et al. [29]	Bangla	Ring radius transform	500	71.91%
Bhattacharya et al. [31]	Devanagri & Bangla	Binarization & morphological operations	100	69.97%

The accuracy of any text detection algorithm depends on the number of text regions (genuine and false) being detected. The false (non-text) regions can be eliminatedby using other features to improve the accuracy. Text and non-text region identification thus becomes a major challenge in the entire text detection process.

The candidate regions generated during the localization step are identified as text or non-text (genuine or false) regions using Stroke Width Transform (SWT). Distance between two parallel lines is considered as stroke width. Uniform width oftext in one word or a complete sentence is an attracting feature and the same has motivated many researchers in past. SWT techniques are broadly categorized into two groups; Gradient based and Distance transform based. In gradient based techniques, edges are derived from the image and the distance between pixels present on two edges following gradient direction is considered as the stroke width. In distance transform based techniques, distance of the foreground object from the nearest edge or boundary is computed using distance measures and the centre value of the distance transform is used to calculate the stroke width. Identifying dark text on light background and vice versa is another challenge in both the techniques. It is

clear from Fig. 4 that, text region has uniform stroke width and non-text region has non-uniform stroke width with strong variation. Hence, SWT becomes a strong feature in detecting text and non-text regions. Fig. 4 also illustrates generalise steps of distance transformation. Distance transform map derived from binary image is called binary distance transform map. Binary distance transform map of dark text on light background and vice versa are presented in Fig. 4. If text is dark and the background is light, then only the derived binary distance transform map can be used to compute stroke width; otherwise binary image inversion is necessary. This demands a technique to accommodate both the scenarios and automatic extractionof stroke width. The motivation behind this work is to address this issue and to provide an automatic approach to compute SWT. The proposed colour-edge based text localization technique is extended further to derive stroke width from theprocessed image. Stroke width computation is based on a binary distance transform map.

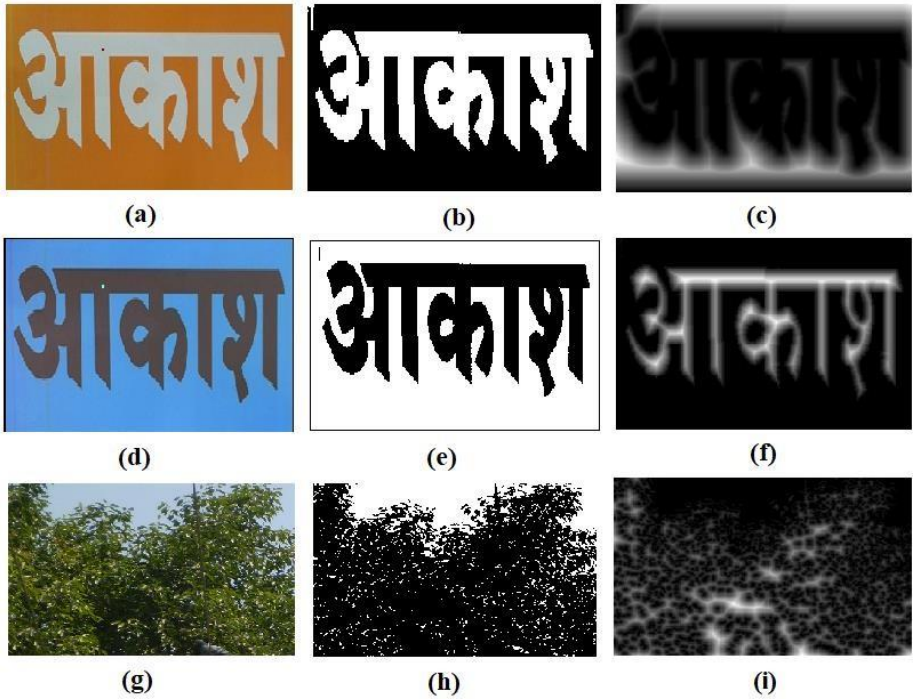


Figure 4. Effect of text and background colour on binary distance transformation map: (a), (b) and (c) are light text on a dark background, its binary version and the corresponding distance transform respectively. (d), (e) and (f)are dark text on a light background, its binary version and the corresponding distance transform respectively. (g), (h) and (i) are non-text region, its binary version and the corresponding distance transform respectively.

Further, SWT along with other geometric features are used for training the classifiers. Different classifiers like SVM, Bayesian and Decision Tree classifier are used for classification of text and non-text regions and the results are compared.750 images are used for training and 500 images are used for testing. Colour-edge based technique as mentioned is applied on images to detect the text regions. The generated regions which match with the ground truth data are considered as text regions and others are considered as non-text regions. If region-based filtering step is omitted, sufficient numbers of non-text regions are obtained for training purpose. A total of 5378 regions are generated by the technique out of which 2757 are non- text regions and 2621 are text regions. All classifiers are trained using appropriate

proportion of the training data set. Classifiers are evaluated by testing total 3736 candidate text regions (1842 text regions and 1894 non-text regions) detected from 500 images by colour and edge-based technique and the performance is measured by deriving a confusion matrix.Table 6 shows the results obtained from Naïve Bayesian classifier. For a detected region, posterior probabilities for text region and non-text regions are computed and the class label with highest probability is selected.

Table 6: Accuracy of Naïve Bayesian classifier

Features	Precision	Recall	F-measure
Geometric features	86.14	76.22	80.88
SWT Features	79.09	66.53	72.27
Geometric + SWT Features	87.43	75.99	**81.31**

Support vector machines are more popular and commonly used classifier for text and non-text region based classification. SVMs' are efficient to handle linear and nonlinear feature sets. It maps original feature set to high dimensional space. Non- linear data is separated using hyperplane in high dimensional space. These hyperplanes are identified with support vectors and margins. Hyperplanes can be computed by taking dot products of support vectors. But, the dot product computations are heavy and costly if used with mapping functions. There are other kernel functions which are resulting into new different non-linear classifiers. The feature set is trained using dot product and radial kernel function and obtained results are mentioned in Table 7.

Table 7: Accuracy of SVM classifier

Features	Operator	Precision	Recall	F-measure
Geometric features	DOT	89.78	72.08	79.96
	Radial	90.1	75.16	81.96
SWT Features	DOT	82.46	68.64	74.92
	Radial	82.03	72.08	76.74
Geometric + SWT Features	DOT	87.9	74.70	80.77
	Radial	93.2	78.01	**84.93**

Decision Tree classifiers are easy to use, highly accurate, handles high dimensional data, does not require prior knowledge base and easy for interpretation. At the time of tree induction, it is very crucial to select splitting feature/attribute as a node at different levels of tree structure. Various attribute selection measures like information gain, gain ratio and gini index are available to decide best attribute for the partitioning. Experimentation is carried out using all the three measures and the results are mentioned in Table 8.

Devanagari script has specific characteristics such as header-lines, upper modifiers and lower modifiers. Similar characteristics are found in Bangla script during the literature survey. Hence, the results of various existing methods for Devanagari and Bangla are considered for the comparison and the comparison is mentioned in Table 9. Different pre-processing techniques and classifiers are used by researchers for text detection. Researchers have used their own dataset which are having less number of images as compared to dataset used here. Compared to the existing

methods, the proposed technique is evaluated on a large dataset of 1250 images containing diversity in font styles like artistic fonts, multi size, multi coloured etc. The dataset used in the proposed method do have diverse characteristics to the existing ones.

Table 8: Accuracy of Decision Tree classifier

Features	Operator	Precision	Recall	F-measure
Geometric features	IG	87.67	78.37	82.76
	GR	88.65	75.30	81.43
	GI	88.74	77.41	82.69
SWT Features	IG	83.21	69.65	75.83
	GR	75.29	79.89	77.52
	GI	83.13	70.80	76.47
Geometric + SWT Features	IG	91.06	79.02	**84.61**
	GR	86.69	78.65	82.47
	GI	90.4	78.28	83.91

Table 9: Comparison of the proposed technique with some of the existing ones.

Techniques	**Script**	**Classifier**	**No. of images**	**Precision**	**Recall**	**F-measure**
Proposed Technique	**Devanagari**	**SVM**	**1250**	**93.2**	**78.01**	**84.93**
T. Kasar et al.	Multiscript (Indian and English)	SVM & NN	100	80.00	86.00	83.00
Aneeshan et al.	Bangla (Horizontal)	HMM	74	82.00	86.00	84.00
	Devanagari (Horizontal)		56	80.00	85.00	82.00
	Bangla (Non- Horizontal)		32	80.00	82.00	81.00
	Devanagari (Non-Horizontal)		43	79.00	83.00	81.00
	Bangla (Curved)		21	74.00	81.00	77.00
	Devanagari (Curved)		32	71.00	77.00	74.00

7.3 Conclusions:

Edge-based technique is capable to handle different colour and size font. Shadow effect and uneven illumination is marginally handled by smoothing technique.Experimental results show that this technique produces high recall but low precision due to irrelevant edge generation. In colour-based technique, clustering is applied on YCbCr and L*a*b* colour spaces. It is observed that YCbCr based technique performs better. Major challenges in text detection like complex background, uneven illumination, shadow effect, multi-size font are tackled using colour-based technique; but multi-colour font is still an issue. Later, the two techniques are combined to improve the accuracy of both the techniques. Irrelevant edge generation in edge-based technique is avoided by applying edge detector on colour-based clustered image. Hence, very few irrelevant edges i.e non-text regions are produced in the combined approach and the overall performance is improved. Also, experimental results show that the proposed technique is more accurate
compared to other existing techniques.

Foreground and background colour-based segmentation reduces generation of irrelevant edges that turn into non-text regions and hence ultimately improves the performance of text detection. Proposed technique is efficient to handle complex background, multi size font, shadow effect, perspective distortion and uneven light conditioning images; but multi coloured font handling is the major limitation. The technique is extended for automatic extraction of stroke width for darker text on lighter background and vice versa. Stroke width is a strong feature for text detectionbut experimental observation states that variance of stroke width is larger forDevanagari script due to the presence of upper and lower modifiers. This restricts the scope of SWT features alone for significant improvement in the results for Devanagari text detection. Experimental results of different classifiers with different operators show that a combination of geometric features and SWT features produces more accurate results. The Decision Tree and SVM classifiers produce nearly same and highly accurate results but it has been observed that SVM with radial function is slightly better than DTC. A large and diversified dataset usedfor the evaluation states that the proposed technique is a better choice for Devanagari text detection compared to existing ones.

7.4 Application Potential:

7.4.1 Long Term

The technology can be transferred to institutes or companies to make sure that the technological development is accessible to a wider range of users who can further develop and exploit the technology into new products, processes, applications, materialsor services. It will also be possible to setup joint ventures and partnerships to share boththe risks and rewards of bringing new technology to market.

7.4.2 Immediate

A camera-based system for cellular phones capable of recognizing and translating Devanagari text from street scene images will be very useful in getting information about route, direction, names, instructions etc. for people from different states and abroad who cannot read Devanagari script.

A data-set containing at least 5000 images in it with various levels of complexity and resolution will allow comparison of various techniques, setting new targets, fostering healthy competition, encouraging collaboration and advancement of the field for the years to come.

8. Research work which remains to be done under the project:
Scene text recognition

Ph.Ds Produced no: Nil	Technical Personnel trained:01	Research Publications arisingout of the present project: Nil (02 papers are under review)

List of Publications from this Project (including title, author(s), journals & year(s)
(A) Papers published only in cited Journals (SCI) : Nil (02 papers are under review)
(B) Papers published in Conference Proceedings, Popular Journals etc.: Nil

Patents filed/ to be filed: Nil

Equipments (Model and Make):

Sr. No	Sanctioned List	Procured (Yes/No) Model & make	Cost (in Rs)	Working (Yes/No)	Utilization Rate (%)
1	Desktop PC (Intel core i5 processor with graphics card and 24 inch monitor)	Yes: DELL OPTIPLEX – 3020, CORE I5 Processor, 4GB DDR3RAM, 500GB H.D.D, DVDWR, Keyboard, Mouse, 24''LED Screen (DELL) 2GB Graphic Card	47,450	Yes	100%
2	Laptop (Intel core i7processor with graphics card)	Yes: DELL LAPTOP-INS -5558, Core i7 Processor(5500U) 8GB RAM , 1TB Hard Disk, 15.6'' Screen, 4GB Graphics Card, Bluetooth, Wi-Fi, Camera,with WINDOWS 8.1	61,900	Yes	100%
3	Laser colour printer	Yes: HP COLOUR LASERJET SINGLE FUNCTION PRINTER –CP-1025	13,850	Yes	50%
4	UPS (1.5 KVA) (Qty: 2)	Yes: APC 1.5 KVA (Qty: 2)	24,700	Yes	100%
5	Cell phone withcamera	Yes: Lenovo K900 16GB, 13MP RearCamera	17,750	Yes	50%
6	All in one (scan, copy, print) blacktoner printer	Yes: HP BLACK AND WHITE LASERJET ALL-IN-ONE PRINTER –M 1136 MFP	9,550	Yes	100%